

XX PA (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
XX PI Hiscott J, Lin R;
XX DR WPI; 1999-620201/53.
XX DR N-PSDB; AAX30993.
XX PT Carboxy-terminus modified highly active forms of interferon regulatory
XX PT factor proteins used for the treatment of viral infections
XX PS Claim 14; Fig 12; 93pp; English.
XX CC The present protein sequence is the interferon regulatory factor-7
XX CC (IRF-7) 2D protein, that is modified in the transactivation domain. The
XX CC serine and threonine residues in the carboxy terminus are modified
XX CC post-translationally, by phosphorylation, following sendai virus
XX CC infection. The modified IRF-7, substituted with aspartic acid, functions
XX CC as a activator of promoters containing ISRE regulatory elements and
XX CC stimulation of DNA binding and transcriptional activity. IRF-7 protein
XX CC shows highest homology to IRF-3. The modified IRF proteins are used for
XX CC the treatment of viral infections like, influenza, herpes or HIV
XX CC infection. They may also be used to activate a cytokine gene, in cancer
XX CC treatment or to modify a target cell of an organism.
XX SQ Sequence 503 AA;
Query Match 100.0%; Score 2731; DB 20; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.6e-197;
Matches 503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTNFRCALESTRFRVWLRDNGSDPADPHK 120
DB 61 KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTNFRCALESTRFRVWLRDNGSDPADPHK 120
QY 121 VVALSRELCHWREGPGTDQTEAEAPAAVPPQGGPPGFLLAHTAGLQAPGLPAPAGDKG 180
DB 121 VVALSRELCHWREGPGTDQTEAEAPAAVPPQGGPPGFLLAHTAGLQAPGLPAPAGDKG 180
QY 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTTSPGPQPAALTGTGEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
DB 241 ETTTSPGPQPAALTGTGEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
QY 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRLRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRLRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARMGKCKVYWEVGGPPGSGASPTACLLPRNCDTPIFDFRVFQSLVEFRAR 420
DB 361 LRGPQLWARMGKCKVYWEVGGPPGSGASPTACLLPRNCDTPIFDFRVFQSLVEFRAR 420
QY 421 QRRGSPRYTILYFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSDLDL 480
DB 421 QRRGSPRYTILYFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSDLDL 480
QY 481 CLSSANSLYDDIECFMELQPA 503
DB 481 CLSSANSLYDDIECFMELQPA 503
RESULT 2
AAE09329
ID AAE09329 standard; Protein; 503 AA.
XX
AC AAE09329;

XX 19-NOV-2001 (first entry)
XX Human intracellular regulatory molecule, KWC02.
DE
XX
XX Human; intracellular regulator; cell division; proliferation; therapy;
KW cancer; infection; wound; developmental abnormality; metabolic problem;
KW cytostatic; antibacterial; vulnery; transcription factor; KWC02;
KW interferon response factor; IRF.
XX
OS Homo sapiens.
XX
XX US6274312-B1.
XX 14-AUG-2001.
XX 10-DEC-1997; 97US-0999774.
XX 11-DEC-1996; 96US-0032818.
XX (SCHE) SCHERING CORP.
PA Gish KC, Seghezzi W, Shanahan F, Lees EM, McClanahan TK;
PI WPI; 2001-535086/59.
XX N-PSDB; AAD16105.
XX
XX New genes encoding intracellular regulatory molecules, useful for
PT regulating cell division and proliferation (e.g. tumor cells),
PT particularly for treating cancer, infections, wounds, or developmental
PT or metabolic abnormalities
XX
XX Example 2; Column 33-38; 62pp; English.
PS
XX The present sequence is a human intracellular regulatory
CC molecule, KWC02. The KWC02, an interferon response factor (IRF)
CC homologue, is a transcription factor. The polynucleotides encoding
CC intracellular regulatory molecules are useful for regulating cell
CC division and proliferation of various cell types, including tumour
CC cells. Specifically, they are also useful for treating cancer,
CC infections, wounds, developmental abnormalities or metabolic problems.
XX
XX Sequence 503 AA;
Query Match 98.7%; Score 2695; DB 22; Length 503;
Best Local Similarity 99.0%; Pred. No. 8.4e-195;
Matches 498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTNFRCALESTRFRVWLRDNGSDPADPHK 120
DB 61 KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTNFRCALESTRFRVWLRDNGSDPADPHK 120
QY 121 VVALSRELCHWREGPGTDQTEAEAPAAVPPQGGPPGFLLAHTAGLQAPGLPAPAGDKG 180
DB 121 VVALSRELCHWREGPGTDQTEAEAPAAVPPQGGPPGFLLAHTAGLQAPGLPAPAGDKG 180
QY 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTTSPGPQPAALTGTGEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
DB 241 ETTTSPGPQPAALTGTGEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
QY 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRLRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRLRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARMGKCKVYWEVGGPPGSGASPTACLLPRNCDTPIFDFRVFQSLVEFRAR 420

```

Db 361 LRGPOLWARMGCKVYVEVGGPGSASPSTPACLLPNCDDTPIDFRVFFOELVEFRAR 420
Qy 421 QRRGSPRYTYILGFGQDLISAGRPKEKSLVLVKLEPWLCRVHLEGTQREGVSSLDSSLDL 480
Db 421 QRRGSPRYTYILGFGQDLISAGRPKEKSLVLVKLEPWLCRVHLEGTQREGVSSLDSSLSL 480
Qy 481 CLSSANSLYDDIECFIMELEQPA 503
Db 481 CLSSANSLYDDIECFIMELEQPA 503

```

RESULT 3

AA15104
ID AAY15104 standard; Protein; 542 AA.

XX AC AAY15104;

XX DT 25-JAN-2000 (first entry)

XX DE Chimeric protein IRF-7 (1-246)/IRF-3 (132-427).

XX KW Interferon regulatory factor; IRF; chimeric protein; serine; threonine;
KW carboxy terminus; amino terminal domain; aspartic acid; phosphorylation;
KW post-translational modification; sendai virus; cancer treatment; herpes;
KW PRDI/PRDIII; promoter; ISRE regulatory element; stimulation; activate;
KW activator; DNA binding; transcriptional activity; viral infection;
KW proteasome mediated degradation; influenza; HIV infection; cytokine gene;
KW target cell.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key

FT Region 1..246

FT /note= "Corresponds to modified amino-terminal domain
of IRF-7 protein"

FT Region 247..542

FT /note= "Corresponds to modified carboxy-terminus of
IRF-3 protein"

FT Misc-difference 511

FT /note= "Wild type Ser replaced with Asp"

FT Misc-difference 513

FT /note= "Wild type Ser replaced with Asp"

FT Misc-difference 517

FT /note= "Wild type Ser replaced with Asp"

FT Misc-difference 519

FT /note= "Wild type Ser replaced with Asp"

FT Misc-difference 520

FT /note= "Wild type Ser replaced with Asp"

XX WO9951737-A1.

XX PD 14-OCT-1999.

XX PF 07-APR-1999; 99WO-CA00314.

XX PR 07-APR-1998; 98CA-2234588.

XX PA (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.

XX PI Hiscott J, Lin R;

XX DR WPI; 1999-620201/53.

XX DR N-PSDB; AAX90994.

XX Carboxy-terminus modified highly active forms of interferon regulatory
PT factor proteins used for the treatment of viral infections
XX Claim 11; Fig 13; 93pp; English.

XX The present sequence is the chimeric protein comprising, residues 1-246
CC from the modified amino-terminal domain of interferon regulatory factor,

CC IRF-7 and 132-427 residues from the carboxy-terminus of modified IRF-3
CC (SD) protein. The serine and threonine residues are post-translationally
CC modified by phosphorylation, following sendai virus infection. The
CC modified IRF, substituted with aspartic acid, functions as a strong
CC activator of promoters containing ISRE and PRDI/PRDIII regulatory
CC elements, stimulation of DNA binding and transcriptional activity and
CC proteasome mediated degradation. The modified IRF sequences are used for
CC the treatment of viral infections like, influenza, herpes or HIV
CC infection. They may also be used to activate a cytokine gene, in cancer
CC treatment or to modify a target cell of an organism.

XX SQ Sequence 542 AA;

Query Match 57.4%; Score 1567.5; DB 20; Length 542;

Best Local Similarity 59.7%; Pred. No. 7.7e-110;

Matches 327; Conservative 38; Mismatches 116; Indels 67; Gaps 12;

Qy 1 MALAPERAAPRVLFGWLLGRISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60

Db 1 MALAPERAAPRVLFGWLLGRISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60

Qy 61 KAWAVARGWPPSSRGSGGPPPEAETAERAGWKTFRCALSTRFRVMLRDNSSGDPADPHK 120

Db 61 KAWAVARGWPPSSRGSGGPPPEAETAERAGWKTFRCALSTRFRVMLRDNSSGDPADPHK 120

Qy 121 VYALSRELWCWREGPGTDTQTEAEAPAAVPPQGGPPGFLATHAGLQAPGPLPAPAGDKG 180

Db 121 VYALSRELWCWREGPGTDTQTEAEAPAAVPPQGGPPGFLATHAGLQAPGPLPAPAGDKG 180

Qy 181 DLLLQAVQOOSCLADHLLTASWGADVPVTKAPGEGOGELPLTGACAGGGLPAGELYGMVAV 240

Db 181 DLLLQAVQOOSCLADHLLTASWGADVPVTKAPGEGOGELPLTGACAGGGLPAGELYGMVAV 240

Qy 241 ETTPTS-----PGPOPAALTTCGAAAPESP HQAEPYLSPS----- 274

Db 241 ETTPTS-----PGPOPAALTTCGAAAPESP HQAEPYLSPS----- 274

Qy 275 -----PSACTAVQEPSG---ALDVTIMYKGRVTLQKVGHPSCTFLYGPDPAPVAT 324

Db 275 -----PSACTAVQEPSG---ALDVTIMYKGRVTLQKVGHPSCTFLYGPDPAPVAT 324

Qy 295 TFPNPLGSENPLKRLVPGSEWEPEVTAFYRGVQVFOQTISCPEGLRLVG-SEVGDRTL 353

Db 295 TFPNPLGSENPLKRLVPGSEWEPEVTAFYRGVQVFOQTISCPEGLRLVG-SEVGDRTL 353

Qy 325 DPQVAFPSPA-ELPDOKQLRYTELLRHVAPGLHLELGPOLWARRMCKKCVYVEVG-- 381

Db 325 DPQVAFPSPA-ELPDOKQLRYTELLRHVAPGLHLELGPOLWARRMCKKCVYVEVG-- 381

Qy 354 PGWPTLPDQMSLTDRGVMSYVRHVLSCGLGGLALWRAGQWLVAQRLGHCHTYVAVSSE 413

Db 354 PGWPTLPDQMSLTDRGVMSYVRHVLSCGLGGLALWRAGQWLVAQRLGHCHTYVAVSSE 413

Qy 382 -----GPGSASPTACLLPNCDDTPIDFRVFFOELVEFRARQRRGSPRYTYILGF 434

Db 382 -----GPGSASPTACLLPNCDDTPIDFRVFFOELVEFRARQRRGSPRYTYILGF 434

Qy 435 GQDLISAGRPKEKSLVLVKLEPWLCRVHLEGTQREGVSSLDSSLDLCLSSANSLSL---YD 491

Db 435 GQDLISAGRPKEKSLVLVKLEPWLCRVHLEGTQREGVSSLDSSLDLCLSSANSLSL---YD 491

Qy 464 GESWPQDQFWTKRLVMVKVPTCLRALVEMARVGGASSLENT-VDLHIDNDHPDLDDDDQ 522

Db 464 GESWPQDQFWTKRLVMVKVPTCLRALVEMARVGGASSLENT-VDLHIDNDHPDLDDDDQ 522

Qy 492 IECFLMEL 499

Db 523 YKAVLQDL 530

RESULT 4

AAW78986

ID AAW78986 standard; Protein; 427 AA.

XX AC AAW78986;

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 1648.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

Db 64 YVP-----GRDKPDLPT-----WKNFRSALNRKEGLRLAEDRSKDPHPDKIYEFV----- 110
QY 130 WREGGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKGDLQLLAQVQQ 189
Db 111 -NSGVG-----DFSQPDTSPTDNG-----GSTSD-----TQ 136
QY 190 SCLADHLITAGWADPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGQ 249
Db 137 EDILDELL-GNMVLAPL-----DFGPP-----SLAVAPEPCQ 169
QY 250 PAALTTGAAAPESPHQAEPLYSPSPSACTAVQEP-SPCALDVTIMYKGRVTLQKVGH 308
Db 170 PURSPLDNPT-----FPNLGSENPLKRLLVPEGEWEFEVTFYRGRQVQTTISCP 223
QY 309 SCTFLYGGPPFAVRATDQQVAPPSPA-ELPDQKQLRYTEELRLHVAPEGLHLELGPQLW 367
Db 224 EGLRLVG-SEVGDRTLPGWVTLDPGMSLTDGVMYSYVRHVLSCGLGGLALWRAGQLW 282
QY 368 ARMKGCKVYEVG-----GPGSASPTPACILPRNCDTPIFPRVFQELVEFR 418
Db 283 AQRLGHCHTYWAVSEELLPSNGHGDGE-----VPKKEGGVFDLGPPIVDLITFT 333
QY 419 ARQREGSPRYTILYFGODLSAGRPKEKSLVLVLEPWL-CRVHLEGTQREGVSSLDSD 478
Db 334 EGSGR-SPRYALWFCVGSWPDQDQWTKRLVMKVVPICLRLVEMARVGGASLENT-V 391
QY 479 DLCLSSANSL-----YDDIECFMEL 499
Db 392 DLHINSHPLSITSDQYKAYLQDL 415
RESULT 6
ID AAY15102 standard; Protein; 427 AA.
XX AAY15102;
XX 25-JAN-2000 (first entry)
XX Modified Interferon Regulatory Factor-3 protein(IRF-3).
KW Interferon regulatory factor; IRF-3; transactivation domain; serine;
KW threonine; carboxy terminus; aspartic acid; mutant; phosphorylation;
KW post-translational modification; sendai virus; cancer treatment; herpes;
KW activator; promoter; PRDI/PRDII; ISRE regulatory element; stimulation;
KW DNA binding; transcriptional activity; transcriptional co-activator;
KW CBP/p300; proteasome mediated degradation; viral infection; influenza;
KW HIV infection; activate; cytokine gene; target cell.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 396 /note= "wild type Ser replaced with Asp"
FT Misc-difference 398 /note= "wild type Ser replaced with Asp"
FT Misc-difference 402 /note= "wild type Ser replaced with Asp"
FT Misc-difference 404 /note= "wild type Ser replaced with Asp"
FT Misc-difference 405 /note= "wild type Ser replaced with Asp"
XX WO951737-A1.
XX 14-OCT-1999.
XX 07-APR-1999; 99WO-CA00314.
XX 07-APR-1998; 98CA-2234588.
XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.

XX Hiscott J, Lin R;
XX WPI; 1999-620201/53.
DR N-PSDB; AAX90992.
XX Carboxy-terminus modified highly active forms of interferon regulatory
XX factor proteins used for the treatment of viral infections -
XX Claim 8; Fig 10; 93pp; English.
XX The present protein sequence is the interferon regulatory factor-3
XX (IRF-3)5D protein, that is modified in the transactivation domain. The
XX serine and threonine residues in the carboxy terminus are modified
XX post-translationally, by phosphorylation, following sendai virus
XX infection. The modified IRF-3, substituted with aspartic acid, functions
XX as a strong activator of promoters containing PRDI/PRDII or ISRE
XX regulatory elements, stimulation of DNA binding and transcriptional
XX activity, association of IRF-3 with the transcriptional co-activator
XX CBP/p300 and proteasome mediated degradation. The modified IRF proteins
XX are used for the treatment of viral infections like, influenza, herpes
XX or HIV infection. They may also be used to activate a cytokine gene, in
XX cancer treatment or to modify a target cell of an organism.
SQ Sequence 427 AA;
Query Match 16.5%; Score 450; DB 20; Length 427;
Best Local Similarity 27.8%; Pred. NO. 8.6e-26;
Matches 140; Conservative 63; Mismatches 193; Indels 108; Gaps 19;
QY 10 PRVLFGWLLGEISSGCGYEGQLWDLDEARTCRVRPKHFKDLSADARIKAWAVARG 69
Db 6 PRIL--PWLVSQDLQGLQEGVAMVWNSKTRPRIPWKHGLRQDAQQEDFGIQAWEATGA 63
QY 70 WPPSSRGGPPPEAETAERAGWKTNFRCALSTRFVMLRNSGDPADPHKVVYALSREL 129
Db 64 YVP-----GRDKPDLPT-----WKNFRSALNRKEGLRLAEDRSKDPHPDKIYEFV----- 110
QY 130 WREGGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKGDLQLLAQVQQ 189
Db 111 -NSGVG-----DFSQPDTSPTDNG-----GSTSD-----TQ 136
QY 190 SCLADHLITAGWADPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGQ 249
Db 137 EDILDELL-GNMVLAPL-----DFGPP-----SLAVAPEPCQ 169
QY 250 PAALTTGAAAPESPHQAEPLYSPSPSACTAVQEP-SPCALDVTIMYKGRVTLQKVGH 308
Db 170 PURSPLDNPT-----FPNLGSENPLKRLLVPEGEWEFEVTFYRGRQVQTTISCP 223
QY 309 SCTFLYGGPPFAVRATDQQVAPPSPA-ELPDQKQLRYTEELRLHVAPEGLHLELGPQLW 367
Db 224 EGLRLVG-SEVGDRTLPGWVTLDPGMSLTDGVMYSYVRHVLSCGLGGLALWRAGQLW 282
QY 368 ARMKGCKVYEVG-----GPGSASPTPACILPRNCDTPIFPRVFQELVEFR 418
Db 283 AQRLGHCHTYWAVSEELLPSNGHGDGE-----VPKKEGGVFDLGPPIVDLITFT 333
QY 419 ARQREGSPRYTILYFGODLSAGRPKEKSLVLVLEPWL-CRVHLEGTQREGVSSLDSD 478
Db 334 EGSGR-SPRYALWFCVGSWPDQDQWTKRLVMKVVPICLRLVEMARVGGASLENT-V 391
QY 479 DLCLSSANSL-----YDDIECFMEL 499
Db 392 DLHIDNDHPLDLDQYKAYLQDL 415
RESULT 7
ID ABP42913 standard; Protein; 175 AA.
XX ABP42913;
XX 25-JAN-2000 (first entry)
XX Modified Interferon Regulatory Factor-3 protein(IRF-3).
KW Interferon regulatory factor; IRF-3; transactivation domain; serine;
KW threonine; carboxy terminus; aspartic acid; mutant; phosphorylation;
KW post-translational modification; sendai virus; cancer treatment; herpes;
KW activator; promoter; PRDI/PRDII; ISRE regulatory element; stimulation;
KW DNA binding; transcriptional activity; transcriptional co-activator;
KW CBP/p300; proteasome mediated degradation; viral infection; influenza;
KW HIV infection; activate; cytokine gene; target cell.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 396 /note= "wild type Ser replaced with Asp"
FT Misc-difference 398 /note= "wild type Ser replaced with Asp"
FT Misc-difference 402 /note= "wild type Ser replaced with Asp"
FT Misc-difference 404 /note= "wild type Ser replaced with Asp"
FT Misc-difference 405 /note= "wild type Ser replaced with Asp"
XX WO951737-A1.
XX 14-OCT-1999.
XX 07-APR-1999; 99WO-CA00314.
XX 07-APR-1998; 98CA-2234588.
XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.

the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. CC Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, CC autoimmune disease or accidental damage. The polypeptides and nucleotides CC may also be used in the diagnosis of the above conditions, and in drug CC screening techniques. The present sequence represents a novel human CC polypeptide of the invention.

XX SQ Sequence 72 AA;

Query Match 12.3%; Score 335.5; DB 22; Length 72;
Best Local Similarity 90.1%; Pred. No. 4.4e-18;
Matches 64; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 152 GGPPGFLATHAGLQAGPLPAPAGDKGDLQLLQVQSCDLADHLTASWGA-DPVPTKA 210
Db 1 GGPPGFLATHAGLQAGPLPAPAGDEGDLALLAVQSCDLADHLTASWGKDIPTKA 60

QY 211 PGEGQGLPLT 221

Db 61 LGEGQGLPLT 71

RESULT 12

ABBI1963
ID ABBI1963 standard; peptide; 392 AA.

XX AC ABBI1963;

DT 11-JAN-2002 (first entry)

DE Human IFN regulatory factor homologue, SEQ ID NO:2333.

Human; cytokine; cell proliferation; cell differentiation; growth factor;
haematopoiesis regulation; tissue growth; immunomodulator; activin;
inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
proliferation; metastasis; cancer; tumour; haematopoietic disorder;
myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
chronic inflammatory condition; proliferative retinopathy;
atherosclerosis; coronary heart disease; arterial ischaemia;
bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; antiinflammatory;
antasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnery; antiulcer.

OS Homo sapiens.

XX WO200157188-A2.

PN 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

PI

XX

DR WPI; 2001-457740/49.

DR N-PSDB; ABA09207.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -

XX Claim 20; Page 287-288; 1963pp; English.

XX Sequences ABB10981-ABBI12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX SQ Sequence 392 AA;

Query Match 11.7%; Score 320; DB 22; Length 392;
Best Local Similarity 26.0%; Pred. No. 4.8e-16;
Matches 119; Conservative 51; Mismatches 176; Indels 112; Gaps 19;

QY 62 ANAVARGWPPSSRGGPPPEAETAEAGWKTFRCALSTRFVLMRNSGDPADPHKV 121
Db 15 AWAETAGAYVP---GRDKPLPT-----WKGNFRSALNRKEGLRLAEDRSKDPHPHKI 65

QY 122 YALSRELWCWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAGPLPAPAGDKGD 181
Db 66 YEFV-----NSVG-----DFSQDTSPTNGG-----GSTSD 93

QY 182 LLLQAVQSCDLADHLTASWGA-DPVPTKA-PGEGQGLPLTGACAGGPGLPAGELYGWAVE 241
Db 94 -----TQEDILDELL-GNVLAPLP-----DPGPP-----SLA 120

QY 242 TTPSPGQPAALTGTGAAAPESPFOAEPYLSPSPSACTAVQEP-SPGALDVTIMYGRTV 300
Db 121 VAPEPCQPLRSPSLDNPTF-----FPNLGSEPNLKELLVPGEWEFEVTAFYGRQV 174

QY 301 LQVYVGHRSCTLYGPPDPAVRATDPQQAFFSPA-ELPDQKQLRYTEELLRHVAPGLHL 359
Db 175 FOOTISCPGGLRVG-SEVGDRTLPGWPTLPDPGMSLTDRGVMSYVRHVLSCIGGLAL 233


```
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.

PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057659.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
XX WPI; 1998-506364/43.
DR N-PSDB; AAV59565.
XX
XX New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 1; Page 573-574; 721pp; English.
XX
XX This sequence represents a secreted human protein encoded by the nucleic
CC acid molecule designated Gene 54 from the human cDNA clone HPRC93
CC (deposited as clone ATCC 97899 and ATCC 209045).
CC The gene can be used to generate fusion proteins by linking to the gene
CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the
CC stability of the fused protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: AAV59511-V59812; amino acid sequences AAW4731-W5026)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 186
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV59511 for described uses).
XX
XX Sequence 373 AA;

Query Match 9.5%; Score 259.5; DB 19; Length 373;
Best Local Similarity 25.2%; Pred. No. 1.6e-11;
Matches 100; Conservative 59; Mismatches 165; Indels 73; Gaps 15;

QY 115 PADPHKVYALSRELWCWREGPGTQTEAEAPAAVPPQGGFFLAHTHAGLQAPGLPA 174
Db 9 PMNPVKIY----QVC-----DIPQPG-----SIINPGSTGS 36
QY 175 PAGDKGDLLOLQVQSQSLADHLITASWGADPVTATAPGSGQGLPL----TGACAGGPG 230
Db 37 APWDEKD---NDVDEDEDEDE-LDQSQHVPITQDTFFPLNINGSMPAPASVGNCS 92
QY 231 PAGELYGWAVETTPSPGPQPAALTGGAAPESPHQAEPLYSPSPACTAVQEPSGALD 290
Db 93 PEAV---W-----PKTEPLEMEVQ---APIQPFYSSPELWISSLPWT-----DLD 132
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:46:08 ; Search time 27 Seconds
(without alignments)
548.138 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAPRVFGWLLG.....SANSLYDDIECFMLELEQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6C COMB.pep:*
6: /cgn2_6/prodata/1/iaa/6D COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2695	98.7	503	US-08-999-774A-10	Sequence 10, Appl
2	454	16.6	427	US-08-705-771-16	Sequence 16, Appl
3	376	13.8	450	US-08-611-280-2	Sequence 2, Appl
4	376	13.8	450	US-09-195-940-2	Sequence 2, Appl
5	376	13.8	450	US-09-562-466-2	Sequence 2, Appl
6	361.5	13.2	450	US-08-611-280-25	Sequence 25, Appl
7	361.5	13.2	450	US-09-195-940-25	Sequence 25, Appl
8	361.5	13.2	450	US-09-562-466-25	Sequence 25, Appl
9	357	13.1	451	US-08-654-482-14	Sequence 14, Appl
10	328.5	12.0	393	US-09-230-371A-29	Sequence 29, Appl
11	306	11.2	425	US-09-230-371A-30	Sequence 30, Appl
12	259.5	9.5	373	US-09-149-476-374	Sequence 374, App
13	228	8.3	107	US-08-654-482-5	Sequence 5, Appl
14	222.5	8.1	108	US-08-654-482-1	Sequence 1, Appl
15	222.5	8.1	108	US-08-654-482-2	Sequence 2, Appl
16	218	8.0	107	US-08-654-482-7	Sequence 7, Appl
17	212	7.8	107	US-08-654-482-6	Sequence 6, Appl
18	166	6.1	108	US-08-654-482-4	Sequence 4, Appl
19	161.5	5.9	108	US-08-654-482-3	Sequence 3, Appl
20	155	5.7	561	US-08-642-255-52	Sequence 52, Appl
21	154.5	5.7	1064	US-08-642-255-62	Sequence 62, Appl
22	153.5	5.6	960	US-09-219-849-5	Sequence 5, Appl
23	152	5.6	355	US-08-483-533-41	Sequence 41, Appl
24	152	5.6	355	US-09-283-471A-41	Sequence 41, Appl
25	152	5.6	355	PCT-US91-06532-3	Sequence 3, Appl
26	148.5	5.4	560	US-07-683-957B-1	Sequence 1, Appl
27	148	5.4	252	US-08-642-255-61	Sequence 61, Appl

28 148 5.4 696 3 US-08-906-865-4 Sequence 4, Appli
29 148 5.4 696 4 US-09-129-868-4 Sequence 4, Appli
30 148 5.4 1057 3 US-08-931-820-4 Sequence 4, Appli
31 148 5.4 1461 4 US-09-585-887-9 Sequence 9, Appli
32 148 5.4 1461 4 US-09-289-578-9 Sequence 9, Appli
33 146 5.3 504 4 US-09-219-849-3 Sequence 3, Appli
34 145 5.3 1384 4 US-08-976-255-11 Sequence 11, Appli
35 144 5.3 777 1 US-08-642-255-53 Sequence 53, Appli
36 143.5 5.3 1057 3 US-08-931-820-1 Sequence 1, Appli
37 143.5 5.3 1341 3 US-08-963-825-18 Sequence 18, Appli
38 143.5 5.3 1341 4 US-09-500-811-18 Sequence 18, Appli
39 143.5 5.3 1341 4 US-09-570-573-18 Sequence 18, Appli
40 143.5 5.3 1341 4 US-09-548-608-18 Sequence 18, Appli
41 140.5 5.1 234 1 US-08-642-255-51 Sequence 51, Appli
42 140 5.1 357 1 US-07-609-716-66 Sequence 66, Appli
43 140 5.1 357 4 US-08-642-255-33 Sequence 33, Appli
44 140 5.1 357 4 US-08-475-411A-66 Sequence 66, Appli
45 140 5.1 357 4 US-08-478-029A-66 Sequence 66, Appli

ALIGNMENTS

RESULT 1
US-08-999-774A-10
; Sequence 10, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: PatentIn Release #1.0, Version #1.30
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,818
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0646
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-774A-10

Query Match 98.7%; Score 2695; DB 4; Length 503;
Best Local Similarity 99.0%; Pred. No. 3e-223; 5; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 5;

```
QY 1 MALAPERAARVLFGEWLLGEISGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAARVLFGEWLLGEISGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGRPPSSRGSGPPPEAEATRAAGWKNFRCAIRSTRFRFVMLRDNSSGDPADPHK 120
DB 61 KAWAVARGRPPSSRGSGPPPEAEATRAAGWKNFRCAIRSTRFRFVMLRDNSSGDPADPHK 120
QY 121 VYALSRELWREGPGTDTOTAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
DB 121 VYALSRELWREGPGTDTOTAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
QY 181 DILLQAVQOQCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGPLPAGELYGWAV 240
DB 181 DILLQAVQOQCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGPLPAGELYGWAV 240
QY 241 ETTSPGQPAALTTGAAAPESPQAEPLSPSPSACTAVQSPGALDVTIMYKRTV 300
DB 241 ETTSPGQPAALTTGAAAPESPQAEPLSPSPSACTAVQSPGALDVTIMYKRTV 300
QY 301 LQKVGHPSCTELYGPPDPVAVRATDPOVAFPPSPAEPLPDQKQLRYTEELRHVAPGLHLE 360
DB 301 LQKVGHPSCTELYGPPDPVAVRATDPOVAFPPSPAEPLPDQKQLRYTEELRHVAPGLHLE 360
QY 361 LRGPQWARMGKCKVYVEGPPGSGASPTPACLLPRNCDTPIFDFRVFQELVEFRAR 420
DB 361 LRGPQWARMGKCKVYVEGPPGSGASPTPACLLPRNCDTPIFDFRVFQELVEFRAR 420
QY 421 QRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 480
DB 421 QRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 480
QY 481 CLSSANSYDDIECFLEMQPA 503
DB 481 CLSSANSYDDIECFLEMQPA 503

RESULT 2
US-08-705-771-16
; Sequence 16, Application US/08705771
; Patent No. 6054289
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; APPLICANT: Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; TITLE OF INVENTION: Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,771
; FILING DATE: August 30, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 427 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-705-771-16

Query Match 16.6%; Score 454; DB 3; Length 427;
Best Local Similarity 28.0%; Pred. No. 4.8e-31;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;

QY 10 PRVLFGEWLLGEISGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIFKAWAVARG 69
DB 6 PRXL--PWLVSQDLGQLEGVAMVWKSRTFRIPWKHGLRQDAQOEDFGIFQAWAETGA 63
QY 70 WPSSRGSGPPPEAEATRAAGWKNFRCAIRSTRFRFVMLRDNSSGDPADPHKVALSREL 129
DB 64 YVP---GRDKPLPT---WKRNFRSALNRKEGLRLAEDRSKDPHPKHYEFV--- 110
QY 130 WREGPGTDTOTAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKGLLQAVQQ 189
DB 111 -NSGVG---DFSQPDTSPDTNGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGPLPAGELYGWAVETTPSPGP 249
DB 137 EDILDELL--GNMVLAPLP-----DPGPP-----SLAVAPEPCPQ 169
QY 250 PAALTGEEAAPPSPHQAEPYLSPPSACTAVQSPGALDVTIMYKRTVLOKVVGHP 308
DB 170 PLRSPSLDNPTP-----FPNLGSPENPLKRLLVPGEEWEFEVTAIFYRGVQFQOTISCP 223
QY 309 SCTFLYGPDPVAVRATDPOVAFPPSPA-ELPDQKQLRYTEELRHVAPGLHLELRGPQLW 367
DB 224 EGLRVG-SEVGRDRTLPGHPVTLDPQMSLTDGVMYSYRVHVLSCGGGLALWRAGOWLW 282
QY 368 ARRMGKCKVYVEVG-----GPPGSGASPTPACLLPRNCDTPIFDFRVFQELVEFR 418
DB 283 AQRLGHCHTYAVSEELLPSNGHGPDE-----VPKDEGGVFDLGPFIVDLITFT 333
QY 419 ARQRCSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 478
DB 334 EGSGR-SPRYALMFCVGSWPDQPPWKRLVMVKKVYVPTCLRALVEMARVGGASSENT-V 391
QY 479 DCLSSANSLS---YDDIECFLEML 499
DB 392 DLHISNSHPLSLTSDQYKAYLQDL 415

RESULT 3
US-08-611-280-2
; Sequence 2, Application US/08611280
; Patent No. 5891666
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,280
; FILING DATE:
```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-611-280-2

Query Match      13.8%; Score 376; DB 2; Length 450;
Best Local Similarity 28.1%; Pred. No. 2.5e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLIGEISGCGYEGLOWLDEARTCFRVWKHPARKDLS-EADARIFKAWAVARGRWPPSS 74
Db 26 QWLIDQIDSGKYPGLVWENEESKVFRIPWKHAGKQDYNREDAALFKAWALFKGKF---- 81
QY 75 RGGGPPPAETAERAGWKNFRCALSTRFRVFLMRDMSG-DPADPHKVVYALSRELWCWREG 133
Db 82 REGIDKPDPT-----WKTRLCALNKSNDFEELVERSOLDISDPKYVTRI----- 127
QY 134 PGTDQTEAEPAAVPPPGGPPGP---FLAHTHAGLOAPGPPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKKGAKQLTLDDTQMAMGHYPMTAP---YGSLLPAQOVHNY 169
QY 191 CLADHLLTASGADVPYTKAPGEGEGLPLTGACAG-----GGLP-----AGELYGWAWE 241
Db 227 ESQAPGPIEPSIRSAEALA-----LSDC-----RLHICLYRDLIV 263
QY 301 LQKVVGHP-SCFTLYGPPDPVAVRATDPOQVAFPPSPAEPLPDQKQLRYTBELLRHVAPGLHL 359
Db 264 KELTTTSPGCRISHG---HTYDVSNLDQVLFYP---DDNGQRKNIEKLLSHLERGLVL 317
QY 418 RARQRGSPRYTYLFGODLSAGRPKEKSLVLKLEPWLCL 459
Db 372 -AHGGRPAFRFQVTLFCFGEFFDPDQ-RORKLITAHVEPLLAR 411

RESULT 4
US-09-195-940-2
; Sequence 2, Application US/09195940
; Patent No. 6258935
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,940
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/611,280
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-195-940-2

Query Match      13.8%; Score 376; DB 4; Length 450;
Best Local Similarity 28.1%; Pred. No. 2.5e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLIGEISGCGYEGLOWLDEARTCFRVWKHPARKDLS-EADARIFKAWAVARGRWPPSS 74
Db 26 QWLIDQIDSGKYPGLVWENEESKVFRIPWKHAGKQDYNREDAALFKAWALFKGKF---- 81
QY 75 RGGGPPPAETAERAGWKNFRCALSTRFRVFLMRDMSG-DPADPHKVVYALSRELWCWREG 133
Db 82 REGIDKPDPT-----WKTRLCALNKSNDFEELVERSOLDISDPKYVTRI----- 127
QY 134 PGTDQTEAEPAAVPPPGGPPGP---FLAHTHAGLOAPGPPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKKGAKQLTLDDTQMAMGHYPMTAP---YGSLLPAQOVHNY 169
QY 191 CLADHLLTASGADVPYTKAPGEGEGLPLTGACAG-----GGLP-----AGELYGWAWE 241
Db 170 MPPHP--DRSW-RDYAPDQSHPEIPYQCPVTFGRGHHWQSPSCNGCQVTFYACAPP 226
QY 242 TTPSPG-PQPAALTGEEAAPSPHQAEPYLSPSACTAVQEPSFGALDVTIMYKGRTV 300
Db 264 KELTTTSPGCRISHG---HTYDVSNLDQVLFYP---DDNGQRKNIEKLLSHLERGLVL 317
QY 301 LQKVVGHP-SCFTLYGPPDPVAVRATDPOQVAFPPSPAEPLPDQKQLRYTBELLRHVAPGLHL 359
Db 264 KELTTTSPGCRISHG---HTYDVSNLDQVLFYP---DDNGQRKNIEKLLSHLERGLVL 317
QY 360 ELRGQPLWARRMGKCKVYWEVGGPPGSGAS--PSTPACLLPRNCPTPIPDFRVFFQELVEF 417
Db 318 WMAPDGLYAKRLCQSRITYD--GPLALCSDRPNK----LERDQCKLFDTQOQFLSELQVF 371
QY 418 RARQRGSPRYTYLFGODLSAGRPKEKSLVLKLEPWLCL 459
Db 372 -AHGGRPAFRFQVTLFCFGEFFDPDQ-RORKLITAHVEPLLAR 411

RESULT 5
US-09-562-466-2
; Sequence 2, Application US/09562466
; Patent No. 6369202
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
```

ZIP: LSN 6JB
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/562,466
FILING DATE: 01-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/195,940
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-338A
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-562-466-2

Query Match 13.8%; Score 376; DB 4; Length 450;
Best Local Similarity 28.1%; Pred. No. 2.5e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;
QY 16 EWLLGEISSCYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARWRPPSS 74
DB 26 QWLIDQIDSGKYPLGVWNEEKSVFRIPWKHAGQDYNREEDAAALFKAWLFKGF----- 81
QY 75 RGGGPPPEAETAEAGWKTFRCALRSTRFRVLMRLDNG--DPADPHKVVYALSRELWREG 133
DB 82 REGIDKDPPT-----WKTRLCALNKSNDFEELVERSOLDISDPYKVYRI----- 127
QY 134 PGTDQTEAEAPAAVPPPGGPGP-----FLAHTAGLQAPGLPAPAGDKGDLALLQAVQOS 190
DB 128 -----VPEGAKKAKQLTLDOTQAMGHYPMTAP-----YGSLEPAQQVHNY 169
QY 191 CLADHLLTASWADPVTKAPGEGQGLPLTGACAG-----GPGLP-----AGELYGWAVE 241
DB 170 MPEH--DRSW-RDYAPDQSHPEIPYOCPTFGPRGHWHQPGSCENCQVGTIFYACAPP 226
QY 242 TTSPSG-PQPAALTGTGEAAAPESPHQAEPLYSFSPSACTAVQEPSPGALDVTIMYKGRV 300
DB 227 ESQAPGPIEPSIRSAEALA-----LSDC-----RLHICLYYRDILV 263
QY 301 LQKVVGHP-SCTFLYGPPDPAVRATDPQVAFSPSAPLDPQKQLRYTEELLRHVAPGLHL 359
DB 264 KELTTSPGECRISHG---HTYDVSNDLQVLPFP-----DNGQKNIKLLSHLERGLV 317
QY 360 ELRGPQLWARMGKCKVYWEVGGPPGSGAS--PSTPACLLPRNCDDTPIFDFRVFFQELVEF 417
DB 318 WMAPDGLYAKLQCSRIYWD--GFLALCSDRPNK-----LERDQCKLFDYQQFLSELQVF 371
QY 418 RARQRRGSPRYTYLGFQDLSAGRPKEKSLVLVKLEPWLCR 459
DB 372 -AHHGRPAFRFQVTLFCFEEFPDPQ-RQRKLITAHVEPLLAR 411

RESULT 6
US-08-611-280-25
Sequence 25, Application US/08611280
Patent No. 5891666
GENERAL INFORMATION:
APPLICANT: Matsuyama, Toshifumi
APPLICANT: Grossman, Alex
APPLICANT: Richardson, Christopher D.
TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES

NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Canada Inc.
STREET: 6733 Mississauga Road, Suite 303
CITY: Mississauga
STATE: Ontario
COUNTRY: Canada
ZIP: LSN 6JB
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,280
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-338A
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-611-280-25

Query Match 13.2%; Score 361.5; DB 2; Length 450;
Best Local Similarity 27.2%; Pred. No. 4.4e-23;
Matches 135; Conservative 70; Mismatches 194; Indels 97; Gaps 22;
QY 16 EWLLGEISSCYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARWRPPSS 74
DB 26 QWLIDQIDSGKYPLGVWNEEKSVFRIPWKHAGQDYNREEDAAALFKAWLFKGF----- 81
QY 75 RGGGPPPEAETAEAGWKTFRCALRSTRFRVLMRLDNG--DPADPHKVVYALSRELWREG 133
DB 82 REGIDKDPPT-----WKTRLCALNKSNDFEELVERSOLDISDPYKVYRIVEGAKK-- 134
QY 134 PGTDQTEAEAP--AAVPPPGGPGPPL-AHTAGLQAPGLPAPAGDKGDLALLQAVQOS 190
DB 135 -GAKQLTLEDPQMSGHPVTMTTPPSLPAQVHNYMPP-----GACAGPGLPAGELYGWAV 240
QY 191 CLADHLLTASWADPVTKAPGEGQGLPLT-----GACAGPGLPAGELYGWAV 240
DB 173 -----LDRSW-RDYVDPQPHPEIPYOCPTFGPRGHWHQPGACENGCCQV-TGTFYACAP 224
QY 241 ETTPSPG-PQPAALTGTGEAAAPESPHQAEPLYSFSPSACTAVQEPSPGALDVTIMYKGR 299
DB 225 PESQAFGVTEPSIRSAEALA-----FSDC-----RLHICLYREIL 261
QY 300 LQKVVGHP-SCTFLYGPPDPAVRATDPQVAFSPSAPLDPQKQLRYTEELLRHVAPGLH 358
DB 262 VKELTTSPGECRISHG---HTYDASNLQVLPFP---EDNGQKNIKLLSHLERGV 315
QY 359 LEURGPQLWARMGKCKVYWEVGGPPGSGAS--PSTPACLLPRNCDDTPIFDFRVFFQELVE 416
DB 316 LMAPDGLYAKLQCSRIYWD--GFLALCNDRENK-----LERDQCKLFDYQQFLSELQ 369
QY 417 FRARQRRGSPRYTYLGFQDLSAGRPKEKSLVLVKLEPWLCRVHLEGTREGVSSLDSS 476
DB 370 F-AHHGRSLPRFQVTLFCFEEFPDPQ-RQRKLITAHVEPLLARQLYFAQNSGHFLRGY 427
QY 477 DLDCLLSSANSLYDDI 492
DB 428 DLPEHISNPEDYHRSI 443

RESULT 7
US-09-195-940-25

```
; Sequence 25, Application US/09195940.
; Patent No. 6258935
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Groszeman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,940
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/611,280
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-195-940-25

Query Match
Best Local Similarity 13.2%; Score 361.5; DB 4; Length 450;
Matches 135; Conservative 70; Mismatches 194; Indels 97; Gaps 22;

QY 16 EWLGEISSGCGYGLQWLDEATCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 QWLIDQIDSGKYPGLVWENEKSI FRIPWKHAGKQDYNREDAALFKAWALFKGKF--- 81

QY 75 RGGGPPPEAETAEAGWKNFRCALRSTRFVWLDRNSG-DPADPHKVALSRELCWREG 133
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 REGIDKPDPPPT-----WKRLCALNKSNDFEELVERSQLDISDPYKYRIVPEGAKK-- 134

QY 134 PCTDOTEAEAP--AAVPPPGGPPGFL-AHTHAGLOAPGLPAPAGDKDLLLLQAVQOS 190
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 -GAKQLTLEDQPMQSMHPTMTTPYPSLPAQVHNYMPP----- 172

QY 191 CLADHLLTASWAGDVPPTKAPGEGGGLPLT-----GACAGGPGLPAGELYGMVAV 240
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 -----LDRSM-RDYVPDPQHPHPIYQCMTFGRGHWHQGPACENGCOV-TGTFVACAP 224

QY 241 ETTTSPG-PQPAALTTGEMAAEPESHQABPYLSPSPSACTAVQEPSGALDVTIMYKGT 299
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 PESQAPGVPTSPSIRSAEALA-----PSDC-----RLHICLYREIL 261

QY 300 VLQKVGH-P-SCTFLYCPDPAVRATDPQVAFPSPAELPDQKQRYTEELRHVAPGLH 358
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 262 VKELTTSPEGCRISHG---HTYDASNLQDVLFPYP---EDNGQRKNIEKLHLERGVV 315

QY 359 LELRGQLWARRMCKVYWEVGGPGSAS--PSTACLLPRNCPTPIFDFRVFFOELVE 416
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 LHWAPGLYAKLCQSRITWD--GPLALCNDRENK-----LERDQTKLFDQTQFSELOA 369

QY 417 FRARQRGRSPRTIYLGFGQDLISAGRPKEKSLVLKLEPWLKRVHLEGTQREGVSLDSS 476
```

; Sequence 25, Application US/09195940.

; Patent No. 6258935

; GENERAL INFORMATION:

; APPLICANT: Matsuyama, Toshifumi

; APPLICANT: Groszeman, Alex

; APPLICANT: Richardson, Christopher D.

; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESS: Amgen Canada Inc.

; STREET: 6733 Mississauga Road, Suite 303

; CITY: Mississauga

; STATE: Ontario

; COUNTRY: Canada

; ZIP: L5N 6J8

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/195,940

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/611,280

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Oleski, Nancy A.

; REGISTRATION NUMBER: 34,688

; REFERENCE/DOCKET NUMBER: A-338A

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 450 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-195-940-25

Query Match

Best Local Similarity 13.2%; Score 361.5; DB 4; Length 450;

Matches 135; Conservative 70; Mismatches 194; Indels 97; Gaps 22;

QY 16 EWLGEISSGCGYGLQWLDEATCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74

; :

Db 26 QWLIDQIDSGKYPGLVWENEKSI FRIPWKHAGKQDYNREDAALFKAWALFKGKF--- 81

QY 75 RGGGPPPEAETAEAGWKNFRCALRSTRFVWLDRNSG-DPADPHKVALSRELCWREG 133

; :

Db 82 REGIDKPDPPPT-----WKRLCALNKSNDFEELVERSQLDISDPYKYRIVPEGAKK-- 134

QY 134 PCTDOTEAEAP--AAVPPPGGPPGFL-AHTHAGLOAPGLPAPAGDKDLLLLQAVQOS 190

; :

Db 135 -GAKQLTLEDQPMQSMHPTMTTPYPSLPAQVHNYMPP----- 172

QY 191 CLADHLLTASWAGDVPPTKAPGEGGGLPLT-----GACAGGPGLPAGELYGMVAV 240

; :

Db 173 -----LDRSM-RDYVPDPQHPHPIYQCMTFGRGHWHQGPACENGCOV-TGTFVACAP 224

QY 241 ETTTSPG-PQPAALTTGEMAAEPESHQABPYLSPSPSACTAVQEPSGALDVTIMYKGT 299

; :

Db 225 PESQAPGVPTSPSIRSAEALA-----PSDC-----RLHICLYREIL 261

QY 300 VLQKVGH-P-SCTFLYCPDPAVRATDPQVAFPSPAELPDQKQRYTEELRHVAPGLH 358

; :

Db 262 VKELTTSPEGCRISHG---HTYDASNLQDVLFPYP---EDNGQRKNIEKLHLERGVV 315

QY 359 LELRGQLWARRMCKVYWEVGGPGSAS--PSTACLLPRNCPTPIFDFRVFFOELVE 416

; :

Db 316 LHWAPGLYAKLCQSRITWD--GPLALCNDRENK-----LERDQTKLFDQTQFSELOA 369

QY 417 FRARQRGRSPRTIYLGFGQDLISAGRPKEKSLVLKLEPWLKRVHLEGTQREGVSLDSS 476

Db 225 PESQAPGVPTPSIRSAEALA-----FSDC-----RLHICLYREIL 261
QY 300 VLQKVVGHP-SCFTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTELLRHVAPGLH 358
Db 262 VKELTTSSPGCRISHG---HTYDASLDQVLPYP---EDNQKRNIEKLLSHLGRGV 315
QY 359 LELRGPOLWARRMGKCKVYWEVGGPGSGAS--FSTPACLLPRNCDTPIDFRVFFOELVE 416
Db 316 LWMADPGLYAKRLCQSIYWD--GPLALCNDPNK---LERDQCKLFDTCQFLSELQA 369
QY 417 FRARQRGSPRYTYLGFQDLSAGRPKEKSLVLVLEPWLCRVHLEGTQREGVSSLDSS 476
Db 370 F-AHHGRSLRPFQVTLFCGEEFPDPQ-RQRKLITAHVEPLLARQLYFAQQNSGHFLRGY 427
QY 477 DLDLCLSSANSLYDDI 492
Db 428 DLPEHISNPEDYHRSI 443
RESULT 9
US-08-654-482-14
; Sequence 14, Application US/08654482
; Patent No. 6245562
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN
; TITLE OF INVENTION: MULTIPLE MYELOMA
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/654,482
; APPLICATION NUMBER: US/08/654,482
; FILING DATE: 28-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 50995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-654-482-14
Query Match 13.1%; Score 357; DB 4; Length 451;
Best Local Similarity 27.3%; Pred. No. 1.1e-22;
Matches 135; Conservative 73; Mismatches 194; Indels 92; Gaps 24;
QY 16 EMLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGWPPSS 74
Db 26 QWLIDQIDSGKYPGLVWENEEKSIFRIPWKHAGQDYNBEDAALFKAWALFKGK---- 81
QY 75 RGGPPPEAETAERAGWKTNFRCALSTRFRVLMRLNSG-DPADPHKVVYALSRELWCWREG 133
Db 82 REGIDKEDPPT-----WKTRLCALNKNSDFEELVERSQLDISDPKYKVIPEGAKK-- 134
QY 134 PGTDQTAEPAAVPPPGGPPGPFLLAHTHAGLOAPGP-LPAPAGDKGDLILLQVQOSCL 192

Db 135 -GAKQLTLEDPOQ-----SMSPYT-MTTPYPSLPA-----QQVHNYMM 171
QY 193 ADHLTASGADVPPTKAPGEGQGLPT-----GACAGGPGLPAGELYGWAVET 242
Db 172 PP--LDRSW-RDYVPQPHPEIPYQCPMTFGPRGHHWQGPACENGCOV-TGTIYACAPPE 227
QY 243 TSPSG-PQPAALTTGSAAPESPHQAEPLSPSPSACTAVQEPSGALDVTIMYKGRTVL 301
Db 228 SQAPGVPTPSIRSAEALA-----FSDC-----RLHICLYREILVK 264
QY 302 QKVVGHP-SCFTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTELLRHVAPGLHLE 360
Db 265 ELTTSSPGCRISHG---HTYDASLDQVLPYPEDNGHRKNI---ENLLSHLGRGVLM 318
QY 361 LRGPOLWARRMGKCKVYWEVGGPGSGAS--PSTPACLLPRNCDTPIDFRVFFOELVEFR 418
Db 319 MAPDGLYAKRLCQSIYWD--GPLALCNDPNK---LERDQCKLFDTCQFLSELQA- 371
QY 419 ARQRGSPRYTYLGFQDLSAGRPKEKSLVLVLEPWLCRVHLEGTQREGVSSLDSSDL 478
Db 372 AHHGRSLRPFQVTLFCGEEFPDPQ-RQRKLITAHVEPLLARQLYFAQQNSGHFLRGYDL 430
QY 479 DLCLSSANSLYDDI 492
Db 431 PEHISNPEDYHRSI 444
RESULT 10
US-09-230-371A-29
; Sequence 29, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-29
Query Match 12.0%; Score 328.5; DB 4; Length 393;
Best Local Similarity 25.7%; Pred. No. 2.5e-20;
Matches 120; Conservative 69; Mismatches 179; Indels 99; Gaps 19;
QY 17 WLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGWPPSSR 75
Db 15 WVVEQVESGQPGVCHDDTAKTFRIPWKHAGQDFREDQDAFFKAWALFKGKY----- 69
QY 76 GGGPPPEAETAERAGWKTNFRCALSTRFRVLMRLNSG-DPADPHKVVYALSRELWCWREG 134
Db 70 -----KEGDTGGPAVWKTRLCALNKSSSEPFKEVPERGRMDVAEPKYVQLLPPGIVSGQP 124
QY 135 GTDQTAEPAAVPPPGGPPGPFLLAHTHAGLOAPGP-LPAPAGDKGDLILLQVQOSCLAD 194
Db 125 GTQKVPK-----ROHSSVS-----SERKEED-----AMQNCITLSP 156
QY 195 HLITASGADVPPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSPGPQPAALT 254
Db 157 SVLQDSLNNEE-----EGASG---GAVHSDIG-----SSSSSSSPPEQVETDT 196
QY 255 TGEAAPAPESPHQAE-----PYLSPSPSACTAVQEPSPG-ALDVTIMYKGRTVLQKVGHPS 309

Db 197 T-----EAPFGDQSRSLFLP-----PEDYSLITFYNGRVGAEQVQSLD 240
Qy 310 CTFLYGPDPVAVRATDPOQVAFPSPAELPDQKRLRYTEBLLRHVAPGLHLELRGPOLMAR 369
Db 241 CRLV---AEPGSESSMEQVLPFKPGLP-----TQRLLSQLERGIILVASNPGFLVQ 291
Qy 370 RMGKCKVYWEVGGPGSASPSFACLLPNCDDTPIPIFRVFFQELVEFRARQRG-SPRY 428
Db 292 RLCPISWNAFQAPPGPH-----LLPSNECVLEFRTAYFCRDLY--RYFQGLGPPPKF 345
Qy 429 TTVLFGQDLSAGRPKEKSLVLKLPWLCRVHLEGT--OREGVSSL 473
Db 346 QVTNLWEESHSSHTPQNLTIVKMEQAFARYLLEQTPEQQAAILSL 392
RESULT 11
US-09-230-371A-30
; Sequence 30, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 30
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-30

Query Match 11-2%; Score 306; DB 4; Length 425;
Best Local Similarity 23.5%; Pred.No. 2.4e-18;
Matches 123; Conservative 66; Mismatches 168; Indels 166; Gaps 22;
Qy 16 EMLGEISSCYEGLOWLDEARTCFVPWKHFARKDLS-EADARIPKAWAVARGRWPPSS 74
Db 12 QMLIEQIDSSMYPLIWEENEKSMFRIPKHACKQDYNQEVDAIFKAWAVFKGK- 67
Qy 75 RGGPPPEAETFRAGWKNFRCALRSTRFRVWLRNDSG-DPADPHKVVYALSREL----- 128
Db 68 -----KEGDKAEPATWTKRLRCLNKSDFBEVTRDSQLDISSEPKYVRIVPEEDQCK 121
Qy 129 -----CWREGTGTDTAEAPAAVPPQGGPGFLAHTAGLQAPGLPAPAGDKGD 181
Db 122 LGVATAGCVNEVTEMECGSEIDELKEPS-----VDDYMGIMTKRSPSPDA----- 168
Qy 182 LLLQAVQOCSLADHLITASGADPVPTKAPGEGQGLPLTGACAGPGLPAGELYGWA 241
Db 169 -----CRSQ--LLPDW-----WAHE 181
Qy 242 TTSPGQPAALTGTGAAPESPHQAEPLYSPSPSACTAVQSPSGALDVTIMYKGRVYL 301
Db 182 --PSTGRR--LVTGVTTY-DAHSAF-----SQWISFYVGGKLVG 217
Qy 302 QKVGHF-SC-----TFLYCPDPBAVRATDPOQVAFPSPAELPDQKRLYTEEL 349
Db 218 QATTTCPCECRISLSQGLPGTKLYGP-----EGLELVRFPFPADTIFESRQVTRKL 270
Qy 350 LRHVAPG-LHLELRGPOLWARMGCKVYWEV-----GGPPGSASBSTACILLPNCDFP 403
Db 271 FGHLERGVLSHRSQG--VFVKLCQGRVFCVVVVVCKGRPNK-----LERDEVVQ 319

Qy 404 IFDRVRFQELVEFRARQRGSPRYTYLGFQODLSAGRPKEKSLVLKLPWLCRVHLE 463
Db 320 VEDTSQIFRELQQFYNSQGR-LPDGRVVLFCFGEFPDMAPRSKLILVQIE-----QLYVR 374
Qy 464 GTOREGVSSLDSSDL-----DLCSSANSLSY 489
Db 375 QLAEEAGKSCGAGSVWQAPPEPPDQVFRMFPDICASHORSFF 417
RESULT 12
US-09-149-476-374
; Sequence 374, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 9.5%; Score 259.5; DB 4; Length 373;
Best Local Similarity 25.2%; Pred. No. 2e-14;
Matches 100; Conservative 59; Mismatches 165; Indels 73; Gaps 15;

Qy 115 PADPHKVYALSRELCHWREGPGTDTAEAPAAVPPGGPPGFLAHTHAGLQAGPLPA 174
Db 9 PMNFVKIY-----QVC-----DIPQPG-----SIINPGSTGS 36

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/654,482
APPLICATION NUMBER: US/08/654,482
FILING DATE: 28-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 50995
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-654-482-2

Query Match 8.1%; Score 222.5; DB 4; Length 108;
Best Local Similarity 41.2%; Pred. No. 5.6e-12;
Matches 47; Conservative 21; Mismatches 35; Indels 11; Gaps 4;
Qy 16 EWLIGEISSGCEGLQWLDEARTCFRVPWKGFARKDLS-EADARIFKAWAVAGRWPPSS 74
Db 4 QWLIDQIDSGKYPGLVWNEEKSVFRIIPWKHAGQDYNREEDAALFKAWALFKGF---- 59
Qy 75 RGGGPPPEAETAEARAGWKTFRCLRSTRFRVLMRDNSG-DPADPHKVVALSRE 127
Db 60 REGIDKDPPT-----WKTRLCALNKSNDFEELVERSOLDISDPYKYRIVPE 108

Search completed: June 18, 2003, 12:48:51
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:47:53 ; Search time 50 Seconds
(without alignments)
1088.561 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSYDDIECFLEMELEQPA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 segs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/prodata/2/pubpaa/FCI_NEW_PUB pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/prodata/2/pubpaa/FCIUS_PUBCOMB pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	454	16.6	427	9	US-09-975-253-2
2	259.5	9.5	373	9	US-09-809-391-374
3	167	6.1	1023	9	US-09-893-519A-14
4	153	5.6	349	10	US-09-919-497-76
5	150.5	5.5	550	9	US-09-976-740-47
6	150.5	5.5	550	12	US-10-023-529-47
7	150.5	5.5	550	12	US-10-023-523-47
8	148.5	5.4	560	9	US-10-162-223-7
9	148.5	5.4	561	9	US-09-813-398-23
10	148	5.4	696	9	US-10-122-805-4
11	148	5.4	1464	9	US-10-060-036-159
12	148	5.4	1464	9	US-10-171-311-36
13	148	5.4	1464	9	US-10-216-705-21
14	148	5.4	1464	9	US-10-149-352-2
15	145.5	5.3	1516	9	US-10-060-036-166
16	143	5.2	1497	9	US-10-157-031-128
17	142	5.2	362	9	US-10-050-882-68
18	142	5.2	394	9	US-10-050-882-137
19	140.5	5.1	1259	9	US-10-260-715-8

20	139	5.1	1207	10	US-09-927-112-2
21	138	5.1	325	10	US-09-927-112-10
22	138	5.1	376	10	US-09-955-518-20
23	137	5.0	538	9	US-09-976-740-43
24	137	5.0	538	12	US-10-023-529-43
25	137	5.0	538	12	US-10-023-523-43
26	137	5.0	714	9	US-10-233-885-44
27	137	5.0	714	9	US-10-231-581-44
28	137	5.0	897	9	US-10-099-895-1
29	136.5	5.0	527	9	US-09-854-133-216
30	136.5	5.0	527	10	US-09-738-973-216
31	136	5.0	638	9	US-10-038-010-4
32	134	4.9	503	9	US-10-078-547-2
33	134	4.9	591	9	US-09-291-417-103
34	134	4.9	699	9	US-10-121-988-143
35	133.5	4.9	731	9	US-10-086-464-17
36	133.5	4.9	1274	9	US-10-020-215-2
37	133	4.9	881	10	US-09-816-860A-2
38	132.5	4.9	417	10	US-09-963-285-4
39	132.5	4.9	501	10	US-09-963-285-2
40	132.5	4.9	501	10	US-09-963-285-10
41	132.5	4.9	507	9	US-10-078-547-24
42	132	4.8	1647	9	US-09-824-574-4
43	131.5	4.8	945	8	US-08-965-272-2
44	131.5	4.8	945	12	US-10-121-882-2
45	131.5	4.8	1339	9	US-10-082-830-282

ALIGNMENTS

RESULT 1

US-09-975-253-2
; Sequence 2, Application US/09975253
; Patent No. US20020164694A1
; GENERAL INFORMATION:
; APPLICANT: Paul Moore et al.
; TITLE OF INVENTION: Methods and Compositions for Treating and Preventing Infection
; FILE REFERENCE: PF196P1
; CURRENT APPLICATION NUMBER: US/09/975,253
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/239,963
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val

US-09-975-253-2

Query Match 16.6%; Score 454; DB 9; Length 427;
Best Local Similarity 28.0%; Pred. No. 2.1e-21;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;

QY	10	PRVLFGWLLGISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVGR	69
DB	6	PRXL--PWLVSQDLQLEGVAVNKSRTFRIPWKHGLRQDAQQDFGIFQAWAETGA	63
QY	70	WPSSRGCGPPPEAETAEAGKTKFRCALRTRFRVLMRLDNGSDPADPHKVYALSREL	129
DB	64	YVP---GDKDPDLPT-----WKNFRSALNKEGURLAEDRSKDPHPHKIYEFV----	110
QY	130	WREGFGTDTQTEAEAPAAVPPQCGPPGPFPLATHAGLQAPGFLPAPAGDKGDLLOAVQ	189
DB	111	-NSGVG-----DFSQPDTSPTDNGG-----GSTD-----TQ	136
QY	190	SCLADHLLTASWAGADPVTKAPGEGQEGPLTGCACAGGPGLPAGELYGNWAVETTPSPGP	249

Db 137 EDILDELL-GNMVLAPLP-----DPGPP-----SLAVAPERCPQ 169
QY 250 PAALTGCEAAPSPOAEPYLSPSACTAVOEP-SPGALDVTIMYKGTVLQKVGH 308
Db 170 PLRSPSLDNTP-----FNLGSENPRLKRLVPGBEWFEFTAFYGRQVFOOTISCP 223
QY 309 SCTFLYPPDPVAVRATDPQOQVAPSPA-ELPDQKQLRYTEELRHVAPGLHLELRGPQLW 367
Db 224 EGLRLVG-SEVGDRTLPGWPTLPDPGMSLTDGVMVSVRHLVSLCGGLALWRAGOWLW 282
QY 368 ARMKGKVKVWEVG-----GPGSASPTPACLLPRNCPTIPDFRVFFQELVEFR 418
Db 283 AQRGLGHCHTYWAVSEELLNPSGHGPDGE-----VPMKKEGVDLPGFFVLDLITFT 333
QY 419 ARQRRGSPRYTILYFGQDLSAGRPKEKSLAVKLEPWLRCRVHLEGTQREGVSSLSDDL 478
Db 334 EGGSR-SPRYALMFCVGSWPQDPWTKRLVMKVVPTCLRALVEMARVGGASSLENT-V 391
QY 479 DLCLSSANSI---YDDIECFMEL 499
Db 392 DLHISNHPLSLSDQYKAYLODL 415

RESULT 2

US-09-809-391-374
; Sequence 374, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 374
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (373)
; OTHER INFORMATION: Xaa equals stop translation
US-09-809-391-374

Query Match 9.5%; Score 259.5; DB 9; Length 373;
Best Local Similarity 25.2%; Pred. No. 3.7e-09;
Matches 100; Conservative 59; Mismatches 165; Indels 73; Gaps 15;
QY 115 PADPHKYALSRELWCWREGPGTDQTEAEAPAAVPPPGQGGPPGPFLLAHTHAGLQAPGLPA 174
Db 9 PMNPVKIY---QVC-----DIPQPG-----SIINPGSTGS 36
QY 175 PAGDKGDLLOAQVQOSCLADHLLTASWGADPVPTKAPGEGQGLPL-----TGACAGPGL 230
Db 37 APWDEKD---NDVDEEDEDE-LDQSQHHVPIQDTFFPLNINGSMPAPASVGNCSVGNCS 92
QY 231 PAGELYGWAVETTPSPGQPAALTGTGAAPESPHQAEPLYLSPSACTAVQSPSCALD 290
Db 93 PEAV---W-----PTELEMEVPQ--APIQFFYSPELWISSLPMT-----DL 132
QY 291 VTIYKGTVLQKV-VGHP-SCTFLYGPDPFAVRATD-----PQQVAFSPALPDQKQ 342
Db 133 IKQYRGKEYGQMTVSNPQGCRLFYGLGMPDQDELFGVXLEQVKFGPGEHINEKQ 192
QY 343 LRTEELLRHVAPGLHLELRGPQLWARMKCKYKVEVGGPPGSGASPSSTACILPRNCDT 402
Db 193 KLFTSKLLDVMDRGLILEVSGHAIYAIRLCQCKVYWSGGPCAPSLVAPN-----LIERQKV 248

QY 403 PIFDRVRFQELVEFRARRRGSPPRYTILYFGQDLSAGRPKEKSLVVKLEPWLRCRVHL 462
Db 249 KLFCLTFLSLLDIAHQKGQIEKOPFEIYLCFGEWPDGKPLERKLLILVQVIVVARMY 308
QY 463 EGTQREGVSLDSSDDLCLSSANSYDDIECFMEL 499
Db 309 EMFSGDFTSRFSGSVRLQISTPD-IKDNIVAQLKQL 344

RESULT 3

US-09-893-519A-14
; Sequence 14, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAQ, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/IG548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Corresponds to SEQ ID NO: 87
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/CAA72189
; DATABASE ENTRY DATE: 1997-06-25
; RELEVANT RESIDUES: (1)...(1023)
US-09-893-519A-14

Query Match 6.1%; Score 167; DB 9; Length 1023;
Best Local Similarity 24.6%; Pred. No. 0.0077;
Matches 102; Conservative 22; Mismatches 131; Indels 160; Gaps 20;

QY 62 AMAVAGRWPSPSRGGPPPEAEATBRAGWKTNFRCLARSTRFRVLMRDNNSGDPADPHKV 121
Db 68 AGAAGAGPAAPAEACAPAAPEPPAGRA-----RPGGGGPORD--- 105
QY 122 YALSRELWCWREGPGTDQ---TEAEAPAAV---PPQG--GPPGPFLLAHTHAGLQAPGL 172
Db 106 -----GPPSPRRLVLPAGPAPPAAKURPPPEGSAGACAPVA---AAAVAGPE 151
QY 173 PAPAG---DKGDLLOAQVQOSCLADHLLTASWGADPVPTKAPGEGQGLP-----LTGAC 224
Db 152 PAPAGAKPAGPAALAA-----RAGPGPGPGPGPGPGKPGAGPAGAAQTLLNSA 200
QY 225 A-----GGPG-----LPAGELYGWAVETTPSPGQPAALTGTGAAAPESP 264
Db 201 ALINSHHAAAPAVSLVNNNGPAALLPLPKPAAAGTGTPTPPFVG-----AAAPPAP 250

```

QY      265  HQAEPYLSPSACTAVOEPSGALDVTIMYKGRVQLKVGHPSCTFLYGPDP----- 318
Db      251  AAPSPPAAPAPAPAAAPPPPPA-----PATLARPPGHPA-----GPPTAAPAVP 296
QY      319  -----PAVRATDPOOV-----AFFSP---AELPDQKQLRYTEEL 349
Db      297  PPAAAGNGGSAGCAAPAPAPAGGPAAGVSGQPGPGAAAAAPAGVKAESP-----KRV 348
QY      350  LRHVAPGLH-LRLRGPQLWARRMGKCKYVWVGPPGSGASPSPTACLLPRNCDDTP 403
Db      349  VQAAPPAQAQTLAASGPASTAASM-----VIGPTMQGALPGSAVPPPPAGPTP 395

RESULT 4
US-09-919-497-76
; Sequence 76, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
; LENGTH: 349
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-919-497-76

```

```

Query Match      5.6%; Score 153; DB 10; Length 349;
Best Local Similarity 34.0%; Pred. No. 0.019;
Matches' 35; Conservative 18; Mismatches 40; Indels 10; Gaps 3;

Qy 11 RVLEGEWLLGEISSGCGEGLWLDEARTCFRPVWKHFARKDLS-EADARIFKAWAVAGR 69
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 5 RMRMRPWLEEQINSNTIPGLKWLNKKEKIFQIPWHAARHGWDVEKDAPLFRNRAIHTGK 64

Qy 70 WPSSRGGGPPPEAETRAERAGKWTNFRCALBSTRFRFVWLNRDNS 112
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 65 HQP-----GVDPKDPDKT-----WKANFRCAAMNSLPDIBEVKDKS 98

```

RESULT 5
US-09-976-740-47
; Sequence 47, Application US/09976740
; Publication No: US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47

```

; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
us-09-976-740-47

Query Match          5.5%; Score 150.5; DB 9; Length 550;
Best Local Similarity 23.4%; Presd. No. 0.044;
Matches 89; Conservative 20; Mismatches 111; Indels 137; Gaps 15;

Qy  2  ALAPERAAPRVLGEWLLGRISSGCVYEGQLWDEARTCFRPVWKHFARKDLSADARIFK 61
Db  20  AAAASSAASPHYQEWILDTI-----DSLRS-----RKARPLD-REICRMVR 60

Qy  62  AWAVARGRWPPSSRGCGPPEAEATAE-----RAGWKTNFRCAL--RSTRFRVMLRDNS 112
Db  61  -----RRHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRGA 107

Qy  113  GDPADPHKVYALSRELCWREGPGTQTEAEAPAAVPPQGGPPGPFLLATHAGLQAQPGPL 172
Db  108  TPAPPPAP-----RGGP-----AAAAAPPTPAPPPP-----PAPV 139

Qy  173  PAPAGDKGDLLOAVQOQSLADHLLTASWGADPVPTKAPGEOGEGUPLTGACAGGPGGLPA 232
Db  140  AAAAA-----APARAPRAAAAAAATAPPSEGPAPG 170

Qy  233  GELYGAVETTPSPGPQPAALTTGEAAAPESPHQABPYLSPSPSACTAVQEPSGALDVT 292
Db  171  PR-----AQRAPPLAAPPPAPAAAPFAGPRRA-----PPPAVAARE-SP----- 214

Qy  293  IMYKGRTVLQKVGHPSCTFLYGPPDPANRATDPQOVAFPSPAELPDQKQ 342
Db  215  -----LPPEPPG---PAPFQOQQQPPPPPPPPQ 241

```

```

RESULT 6
US-10-023-529-47
; Sequence 47, Application US/10023529
; Patent No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-529-47

; Query Match          5.5%   Score 150.5;   DB 12;   Length 550;
; Best Local Similarity 23.4%;   Pred. No. 0.044;
; Matches      82;   Conservative      20;   Mismatches      11;   Indels      137;   Gaps
;
; 2 ALAPERAAPRVLFGEWLLGSSCCYEGLOWLDEARCTCFRVVWKFHARKOLSEADARIF

```

Db 20 AAAASSAASPHYQEWILDTI-----DSLRS-----RKARPD-ERICMVR 60
Qy 62 AWAVARGWPPSSRGSGPPPEAETAE-----RAGWKTNFRCAL--RSTRFVMLRDNS 112
Db 61 -----RRHGPEERTAELEKLIQRAVLRSYKGSISYRNAARVQPPRGA 107
Qy 113 GDPADPHKVALSRELWCWREGPGTDQTEAEAPAAVPPPGGPGPFLLAHTHAGLQAPGL 172
Db 108 TPAPPAPRAP-----RGSP-----AAAAAAPPPTAPPPP-----PAPV 139
Qy 173 PAPAGDKDILLQAVQOSCLADHLLTASWGADPVFTKAPGEGQGLPLTGACAGGGLPA 232
Db 140 AAAA-----APARAPRAAAAAAATAAPPSPGPAQPG 170
Qy 233 GELYGWAVETTPSGPQPAALTTGEAAAPESPHQAEPLYSPSPSACTAVQEPSGALDVT 292
Db 171 PR-----AQRAAPLAAPPAPAAAPAGPRA-----PPAAVAARE-SP-----214
Qy 293 IMYKGRVTLQKVGHPSCTFLYGPDPFAVRATDPOQVAFPSPAELPDQKQ 342
Db 215 -----LPPPPQ-----PAPPQQQQPPPPPPPPQ 241

RESULT 7

US-10-023-523-47
; Sequence 47, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-523-47

Query Match 5.5%; Score 150.5; DB 12; Length 550;
Best Local Similarity 23.4%; Pred. No. 0.044;
Matches 82; Conservative 20; Mismatches 111; Indels 137; Gaps 15;

Qy 2 ALAPERAAPRVLFGEWLLGSISSGCVYEGLOWLDEARTCFRVPWKHFARKDLSADARIFX 61
Db 20 AAAASSAASPHYQEWILDTI-----DSLRS-----RKARPD-ERICMVR 60
Qy 62 AWAVARGWPPSSRGSGPPPEAETAE-----RAGWKTNFRCAL--RSTRFVMLRDNS 112
Db 61 -----RRHGPEERTAELEKLIQRAVLRSYKGSISYRNAARVQPPRGA 107
Qy 113 GDPADPHKVALSRELWCWREGPGTDQTEAEAPAAVPPPGGPGPFLLAHTHAGLQAPGL 172
Db 108 TPAPPAPRAP-----RGSP-----AAAAAAPPPTAPPPP-----PAPV 139
Qy 173 PAPAGDKDILLQAVQOSCLADHLLTASWGADPVFTKAPGEGQGLPLTGACAGGGLPA 232

Db 140 AAAA-----APARAPRAAAAAAATAAPPSPGPAQPG 170
Qy 233 GELYGWAVETTPSGPQPAALTTGEAAAPESPHQAEPLYSPSPSACTAVQEPSGALDVT 292
Db 171 PR-----AQRAAPLAAPPAPAAAPAGPRA-----PPAAVAARE-SP-----214
Qy 293 IMYKGRVTLQKVGHPSCTFLYGPDPFAVRATDPOQVAFPSPAELPDQKQ 342
Db 215 -----LPPPPQ-----PAPPQQQQPPPPPPPPQ 241

RESULT 8

US-10-162-223-7
; Sequence 7, Application US/10162223
; Publication No. US20030033637A1
; GENERAL INFORMATION:
; APPLICANT: OISHI, KAREN K.
; APPLICANT: COMARATTA, LEONARD M.
; TITLE OF INVENTION: GENE EXPRESSION AND PRODUCTION OF TGF-B PROTEINS
; TITLE OF INVENTION: INCLUDING BIOACTIVE MULLERIAN INHIBITING SUBSTANCE FROM
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 048281/0140
; CURRENT APPLICATION NUMBER: US/10/162,223
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-223-7

Query Match 5.4%; Score 148.5; DB 9; Length 560;
Best Local Similarity 23.7%; Pred. No. 0.06;
Matches 106; Conservative 24; Mismatches 150; Indels 167; Gaps 21;

Qy 64 AVARGWPPSSRGSGPPPEAETAEARAGWKTNFRCALRSTRF-VMLRDNSSGDPADPHKV 122
Db 86 AVQARWGP--RDLATFGVNTGDR-----QAALPSLRLLGAWLR---DFGGQRLVV 132
Qy 123 ALSRELWCWREGPGTDQTEAEAPAAVPPPGG--GP-----PGFLAHTHAGLQAPG 170
Db 133 LHLEEVWTEPTPSLRPQE-----PPGGAGPPELALLVLPGPGEVTVTRAGL--PG 183
Qy 171 -----PLPAPAGD-KGDLLOAVQOSCLADHLLTASWG-----203
Db 184 AQSCLPSRDTRYLVLAVDPAWAGSGGLALTLPQREGDSRLSTARKLQALLFGDDHRCFT 243
Qy 204 -----DPVPTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGPQPA 251
Db 244 RMTPALLLPSEPAFLPAHQ-----LDTVPFPPRPS 277
Qy 252 ALTTGAAAAPESPHQAEPLYSPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGHPSCT 311
Db 278 -----AELESPSADPFLETLTRLVRALRVP-----PARASAPRLADPDA- 319
Qy 312 FLYGPDPFAVRATDPOV-----APPSPAELPDQKOLRYTEELRH 352
Db 320 -LAGFPQGLVNLSDPAALERLLDGEPEPLLLLRRTAATTGDPAPLHDPTSPATWALARR 378
Qy 353 VAPGLH---LELRG-----PQLWARRMGKCKVYWEVGGPGSPSPACILPRNCD 401
Db 379 VAAELQAAAELSLPGLPPATAPLARLLALCP-----GGPGLGLPLRALLLKA---430
Qy 402 TPFDPRVTFQEL-VEFRARQRRGSPR 427
Db 431 -----LQGLRVEWRGDRPRGCR 448

RESULT 9

US-09-813-398-23
; Sequence 23, Application US/09813398

; CURRENT FILING DATE: 2002-04-15 APPLICATION NUMBER: 09/129,668
 ; PRIOR APPLICATION NUMBER: EARLIER
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-05

Db 489 -----GPCSRGFPAGDVGAGPKG-PAGE----- 510
QY 241 ETTSPGQPAALTTGEAAP-----ESPHQAEPLSPSPSACTAVQEPSGA 288
Db 511 --RGSPGAPGKSPGAGRPGAGLPGAKGLTGSPG-----SPGPDGKTG--PPGPAG 560
QY 289 LDVTIMYKGRVTLQKVGHSPCTFLYGPDPVAVRATDPQOVAFPSPAELPDQQLRYTEE 348
Db 561 QD-----GRP-----GPPGPPGARGQAGVMGPPGPKGAAGE----- 591
QY 349 LLRHVAPGLHLELRGPQLWARRMGKCKVYWEVGGPPGSPASPS 390
Db 592 -----PGKAGE-RG-----VPGPPGAVGPA 610
RESULT 12
US-10-171-311-36
; Sequence 36, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-36

Query Match 5.4%; Score 148; DB 9; Length 1464;
Best Local Similarity 23.4%; Pred. No. 0.18;
Matches 80; Conservative 11; Mismatches 89; Indels 162; Gaps 17;
QY 66 ARGWPPSSRGSGPPPEAETAEAPAAVPPQG--GPPGFLAHTHAGLQA-PGP--LPAPAGDKG 125
Db 414 ARGSPGQPGGPPGP-----KGNSEGEFGAP----- 439
QY 126 RELCWREGPGCTDQTEAEAPAAVPPQG--GPPGFLAHTHAGLQA-PGP--LPAPAGDKG 180
Db 440 -----GSKGDTGAKGE-----PGPVGVQPPGAGEGKRGARGEPGPTGLPGPPGERG 488
QY 181 DLLLQAVQOCLADHLLTASWGADVPVTKAPGSGQEGPLTGACAGPGPLPAGELYGWAV 240
Db 489 -----GPCSRGFPAGDVGAGPKG-PAGE----- 510
QY 241 ETTSPGQPAALTTGEAAP-----ESPHQAEPLSPSPSACTAVQEPSGA 288
Db 511 --RGSPGAPGKSPGAGRPGAGLPGAKGLTGSPG-----SPGPDGKTG--PPGPAG 560
QY 289 LDVTIMYKGRVTLQKVGHSPCTFLYGPDPVAVRATDPQOVAFPSPAELPDQQLRYTEE 348
Db 561 QD-----GRP-----GPPGPPGARGQAGVMGPPGPKGAAGE----- 591
QY 349 LLRHVAPGLHLELRGPQLWARRMGKCKVYWEVGGPPGSPASPS 390

Db 592 -----PGKAGE-RG-----VPGPPGAVGPA 610
RESULT 13
US-10-216-705-21
; Sequence 21, Application US/10216705
; Publication No. US20030096973A1
; GENERAL INFORMATION:
; APPLICANT: Meristem Therapeutics, S.A.
; TITLE OF INVENTION: Recombinant Collagens and Derived Proteins Produced by Plants, Mei
; FILE REFERENCE: 1149-3 DIV
; CURRENT APPLICATION NUMBER: US/10/216,705
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 09/331,347
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-705-21

Query Match 5.4%; Score 148; DB 9; Length 1464;
Best Local Similarity 23.4%; Pred. No. 0.18;
Matches 80; Conservative 11; Mismatches 89; Indels 162; Gaps 17;
QY 66 ARGWPPSSRGSGPPPEAETAEAPAAVPPQG--GPPGFLAHTHAGLQA-PGP--LPAPAGDKG 125
Db 414 ARGSPGQPGGPPGP-----KGNSEGEFGAP----- 439
QY 126 RELCWREGPGCTDQTEAEAPAAVPPQG--GPPGFLAHTHAGLQA-PGP--LPAPAGDKG 180
Db 440 -----GSKGDTGAKGE-----PGPVGVQPPGAGEGKRGARGEPGPTGLPGPPGERG 488
QY 181 DLLLQAVQOCLADHLLTASWGADVPVTKAPGSGQEGPLTGACAGPGPLPAGELYGWAV 240
Db 489 -----GPCSRGFPAGDVGAGPKG-PAGE----- 510
QY 241 ETTSPGQPAALTTGEAAP-----ESPHQAEPLSPSPSACTAVQEPSGA 288
Db 511 --RGSPGAPGKSPGAGRPGAGLPGAKGLTGSPG-----SPGPDGKTG--PPGPAG 560
QY 289 LDVTIMYKGRVTLQKVGHSPCTFLYGPDPVAVRATDPQOVAFPSPAELPDQQLRYTEE 348
Db 561 QD-----GRP-----GPPGPPGARGQAGVMGPPGPKGAAGE----- 591
QY 349 LLRHVAPGLHLELRGPQLWARRMGKCKVYWEVGGPPGSPASPS 390
Db 592 -----PGKAGE-RG-----VPGPPGAVGPA 610

RESULT 14
US-10-149-352-2
; Sequence 2, Application US/10149352
; Publication No. US20030105050A1
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajinder
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 06275-254US1
; CURRENT APPLICATION NUMBER: US/10/149,352
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/GB00/04741
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: GB 9929487.8
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:45:13 ; Search time 22 Seconds
(without alignments)
2197.982 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSLYDDIECFMLEQPA 503
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	780	28.6	491	2 S56753	interferon regulat
2	482.5	17.7	504	2 G02474	interferon regulat
3	440	16.1	459	2 JC6520	interferon regulat
4	376	13.8	450	2 S57837	lymphoid-specific
5	333.5	12.2	424	2 A35861	interferon consens
6	328.5	12.0	393	2 A45017	transcription fact
7	306	11.2	425	2 A45064	interferon consens
8	259.5	9.5	399	2 JC4592	transcription fact
9	224	8.2	203	2 S57836	lymphoid-specific
10	183	6.7	1334	2 T50568	probable multi-dom
11	176	6.4	325	2 I52998	interferon regulat
12	176	6.4	1146	2 A38587	collagen, cornea-s
13	167	6.1	349	2 A53340	interferon regulat
14	166	6.1	325	2 B31595	interferon regulat
15	166	6.1	329	2 A31595	interferon regulat
16	165.5	6.1	328	2 A36330	interferon regulat
17	160.5	5.9	1400	2 T31555	hypothetical prote
18	160	5.9	1747	1 A45374	collagen alpha 1(X
19	159.5	5.8	416	1 SKXLAG	dermal gland prote
20	159	5.8	576	2 T36729	probable serine/th
21	158.5	5.8	839	2 T04859	extensin homolog F
22	157.5	5.8	705	2 A35363	synapsin I splice
23	156.5	5.7	1464	2 S59856	collagen alpha 1(I
24	156	5.7	381	2 T27806	hypothetical prote
25	156	5.7	1857	2 S31212	collagen alpha 1(X
26	156	5.7	1888	2 S78476	collagen alpha 1(X
27	156	5.7	3149	1 Q0888	BPLF1 protein - hu
28	155	5.7	801	2 T29018	hypothetical prote
29	154	5.6	660	1 Q0853	BHLF1 protein - hu

RESULT 1
S56753
interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S56753
R:Grant, C.B.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995
A:Title: CIRP-3, a new member of the interferon regulatory factor (IRF) family that is r
A:Reference number: S56753; MUID:95334365; PMD:7541908
A:Accession: S56753
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <GRA>
A:Cross-references: EMBL:U20338; NID:G790580; PIDN:AAA86995.1; PID:G790581
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match	28.6%;	Score	780;	DB	2;	Length	491;
Best Local Similarity	38.5%;	Pred. No.	3.5e-40;				
Matches	208;	Conservative	60;	Mismatches	180;	Indels	92;
Gaps	17;						
QY	2	ALAPERAPRVLFGEWLLGEISGCVGELOWLDEARTCFRPWKHFARKOLSEADARIFK	61				
DB	3	ALDSEGAQKLRFGPMLLNAVSSGLYRGLCWIDPPRRIFRIPKHNARKDVTSDDVEIFK	62				
QY	62	ANAVARGRWPPSSRGSGPPPEAETAEACGKTNFCALRSTRFRVMLRNSGDPADPHKV	121				
DB	63	AWAKASGRY-----EGNAEDPAKWKTNFCALRSTHMFMLLEDRSKCNDDPHKV	111				
QY	122	YALSRELWCWREGFGTDTQTEAEAPAVPPQP--GGP-----PGPFLAHTHAGLQ----	167				
DB	112	YAVA-----SGVPNDRSGGVPAGALQQQPOLLNHHDLALENTPT	152				
QY	168	-APGPLPAPAGKGL-LLOAVQOCSLADHLITASWADPVPYKAPGEG--QEGLEPLTG-	222				
DB	153	DSTEGVAAALATQVDLDLQSVLQHCNISAL-----GSQFTLWAHTGDALPEDALLPGQ	207				
QY	223	-ACAGGGLPAGELYCM-AVEITPSPGPQP-----AALTGTGAAAPESPQOAE	269				
DB	208	DGCLFGP-----QFQDWRQLEPPLLGNQPLTGGCGGCGAGALPVSECAIAPSPABE	262				
QY	270	YL-----SPSPSACTAVQEPSGALDVTIMYKGRITVLQKVGVHPSCTFLXGPPDPVAVRATD	325				
DB	263	LLFQSANPAPPAGDIGGLPLLDITIIYRGWVYQEQVDDSRCLVAYQLDPAV--AE	320				
QY	326	POQVAPPSPAELPDQKQLRTYTELLRHVAPGLHLEURGPQLWARRMGKCKVYHEV-----	380				
DB	321	QRLVLFPSPASLPDRQRYTELDLE--VAGLRLEQRAGQLLATRLKCKVFWALSQQLE	378				
QY	381	CGPPGASSTPACLLPRNCDTPIEDFRVFFQELVEFRARORGRSPRYTIYLGFGGDLA	440				
DB	379	GGEP-----PUNLHRDQETTFIDFRVFCTELDRDSDRRSPDFTLFLFCGQCFS	431				

unconventional myo
synapsin I - rat
synapsin Ia - rat
hypothetical prote
mullerian inhibiti
hypothetical prote
collagen alpha 1(I
collagen alpha 1(I
hypothetical prote
collagen alpha 1(X
collagen alpha 1(I
collagen 1 - Caeno
KIAA0641 protein -
eyelid - fruit fly
collagen alpha 1(X
collagen alpha 2(X

ALIGNMENTS

C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription regulation

Query Match 12.2%; Score 333.5; DB 2; Length 424;
Best Local Similarity 24.3%; Pred. No. 3.4e-13;
Matches 121; Conservative 65; Mismatches 152; Indels 159; Gaps 19;

Qy 16 EWLLGEISSGCVGEGLOWLDEARTCFRVPWKHPARKDLS-EADARIFKAWAVARGWPPSS 74
Db 12 QWLIQIEDSSMYPGLIWEENDEKTFRIPWKHAGQDYNQEVDDASIFKAWVFKGKF----- 67
Qy 75 RGGGPPPEAETABEAGWKTFRFCALRSTRFRVMLRDNSG-DPADPHKYVALSRELWCWREG 133
Db 68 -----KEGDAEAPATWTKRLURCALNKSPPDEEVTRSQLDISEPKYTRI----- 112
Qy 134 PGTDQTEAEAPAAVPPPPQGGPPGPFLLATHAGLAQAPGLPAPAGDKGDLQLQAVQOOSCLA 193
Db 113 --VPBEEQCKKLG-----APAG-----CWS 131
Qy 194 DHLLTASWADPVPTKAPBGQ-EGLPVTGACAGGGLPAGELYGWAVETTPSPGPQPA 252
Db 132 E-----VPEMECGRSEIBELIKE-----PSVDEYMGMTKRSP----- 165
Qy 253 LTTGAAAPESPQAEPLSPSPACTA-----VQEPSG-----ALD----- 290
Db 166 -----PEACRSQILPDWVQPSAGLPLVTGYAAYDTHHSAPS 203
Qy 291 ---VTIMYKGRTVLQKV-----VGHPSCFTLYGPPDPVAVRATPDQQAFFSPABL 337
Db 204 QMWISFYGKGLVQAYTTCLGCELSLSQPLKLYGPD-----GLEP--VCFTADTI 256
Qy 338 PDQQLRYTEELRHVAPGLHLEURGQPLWARRMKCKVYWEVGGPPGSPASPTPACLLP 397
Db 257 PSEQRQVTRKLFGLHLERGVLLHNRKGVFKRLCQGRVFCGNAVAVCKGRNK-----LE 312
Qy 398 RNCDDTPFDPRVFOELVEFRARQGRSPRTIYLGFGQDLSAGRPKESLVLVKLEPWL 457
Db 313 RDEVVQVFDNQFTRELQOQFYATQSR-LPDSRVVLCFGESEFPDTPVPLRSKLLVQVEQLY 371
Qy 458 CRVHLEGTQRE-GVSSL 473
Db 372 ARQLVEAGKSCGAGSL 388

RESULT 6
A45017
transcription factor ISGF3 gamma chain - human
N;Alternate names: alpha-interferon-responsive transcription factor
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Sep-1999
C;Accession: A45017
R;Veals, S.A.; Schindler, C.; Leonard, D.; Fu, X.Y.; Aebersold, R.; Darnell Jr., J.E.; Le
Mol Cell Biol 12, 3315-3324, 1992
A;Title: Subunit of an alpha-interferon-responsive transcription factor is related to int
A;Reference number: A45017; MUID:92334329; PMID:1630447
A;Accession: A45017
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-393 <VEA>
A;Cross-references: GB:M87503; NID:g184652; PIDN:AAA58687.1; PID:g184653
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:108663)
A;Note: part of this sequence was confirmed by protein sequencing
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription factor

Query Match 12.0%; Score 328.5; DB 2; Length 393;
Best Local Similarity 25.7%; Pred. No. 6.2e-13;
Matches 120; Conservative 69; Mismatches 179; Indels 99; Gaps 19;

Qy 17 WLLGEISSGCVGEGLOWLDEARTCFRVPWKHPARKDLS-EADARIFKAWAVARGWPPSSR 75
Db 15 WVVEQVESGQPPGVCDWDTAKTFRIPWKHAGKQDFREDQDDAAFFKAWAIFKGY----- 69

```
QY 76 GGGPPPEAETAGWKTFRFCALRSTRFVMLRDNSSG-DPADPHKYVALSRELWCWREGP 134
Db 70 -----KEGDTGGPAVWKTRLCALNKSSEFKEVPERGRMDVABEYKYQLLPGLVSGQP 124
QY 135 GTQDTEAEAPAAVPPPGGPPGLFATHAGLQAPGLPAPAGDKGLLQAVQOOSCLAD 194
Db 125 GTQVPSK-----RQHSVS-----SERKEED-----AMQNCTLSP 156
QY 195 HLLTASGADVPVTKAPGEGQGLPLTCACAGGGLPAGELYGWAVETTPSPGPAAALT 254
Db 157 SVLQDSLNNEE-----EGASG-----GAVHSDIG-----SSSSSSSPFQEVTD 196
QY 255 TGEAAAPESPHQAE-----PYLSPSPSACTAVQEPSG-ALDVTIMYKGRVLOKVGHPS 309
Db 197 T-----EAFQGDQDSLEFLP-----PEPYSLLTFIYNGRVVGEAQVQSLD 240
QY 310 CTFLYGPDPFAVRATDPOQVAFPSPAELPQKQLRYTELLRHVAPGLHLELRGPQLWAR 369
Db 241 CRLV---AEPSSGESSMEQVLPKPGPLEP-----TORLLSQLERGILVASNPRGLFVQ 291
QY 370 RMGCKKYVWVGPPGPGSASTPACLLPRNCDTPIPDFRVFFQELVEFRARQRRG-SPRY 428
Db 292 RLCPPIPSWNAPOAPPQPGPH-----LLPSNECVELFTAYFCRDLV--RYFQGLGPPPKF 345
QY 429 TIVLGFQDLSAGRPKEKSLVVKLEPWLCKRVHLEGT--OREGVSSL 473
Db 346 QVTLNFWEEHSGSHSTQNLITVKMEQAFARYLLEQTPEQQAALLSL 392

RESULT 7
A45064
interferon consensus sequence binding protein, ICSBP (DNA-binding domain) - human
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998
C;Accession: A45064
R;Weisz, A.; Marx, P.; Sharf, R.; Appella, E.; Driggers, P.H.; Ozato, K.; Levi, B.Z.
J. Biol. Chem. 267, 25589-25596, 1992
A;Title: Human interferon consensus sequence binding protein is a negative regulator of
A;Reference number: A45064; MUID:93094284; PMID:1460054
A;Accession: A45064
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-425 <WEI>
A;Experimental source: lung and blood
A;Note: sequence extracted from NCBI backbone (NCBI:P:120312)
A;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription regulation

Query Match 11.2%; Score 306; DB 2; Length 425;
Best Local Similarity 23.4%; Pred. No. 1.6e-11;
Matches 121; Conservative 67; Mismatches 174; Indels 156; Gaps 21;

QY 16 EWLGEISSGCEGLQWLDEARTCFRVPWKHFARKDLS-BADARIFKAWAVARGRWPPSS 74
Db 12 QWLIEQIDSSMYPGLIWENEKSMFRIPWKHAGQDVQNEVDASIFKAWAVFKGP---- 67
QY 75 RGGPPPEAETABAGHKTFRFCALRSTRFVMLRDNSSG-DPADPHKYVALSREL----- 128
Db 68 -----KEGDKAEPATWKTRLCALNKSDFEEVTDRLSDISPEYKYRVIPVEEDQCK 121
QY 129 -----CWREGPQDTEAEAPAAVPPPGGPPGLFATHAGLQAPGLPAPAGDKGD 181
Db 122 LGVATACVNEVTEMECGREIDELIKEPS-----VDDYGMIKRSPDPDA----- 168
QY 182 LLLQAVOOSCLADHLLTASGADVPVTKAPGEGQGLPLTCACAGGGLPAGELYGWAVE 241
Db 169 -----CRSQ--LLPDW-----WAHE 181
QY 242 TTSPGPPQPAALTGEAAAPESPHQAEPYLSPSPSACTAVQEPSPGALDVTIMYKGRVTL 301
Db 182 --PSTGRR--LVTGYTTY-DAAHSAP-----SQMVISFYGGKLVG 217
```

```
QY 302 QKVVGHP-SC-----TFLYGPDPFAVRATDPOQVAFPSPAELPQKQLRYTEEL 349
Db 218 QATTTCPEGCRLSLSQGLPGTKLYGP-----EGLELVRFPADTIPSERQROVTRKL 270
QY 350 LRHVAFG-LHLEURGOLWARRMGKVKYVWVGPPGSGASPSPTACLLPRNCDTPIPDFR 408
Db 271 FGLERGVLHSSRQG--VFVKRLCOGRVFCGNAVVCGRPNK-----LERDEVVQVFDTS 324
QY 409 VFFQELVEFRARORRGSPRYTIVLGFQDLSAGRPKEKSLVVKLEPWLCKRVHLEGTQRE 468
Db 325 QFFRELQOQFYNSQGR-LPDGRVVLFCGEEPPDMAPLRSLKILVQIE-----OLYVRQLAEE 379
QY 469 GVSLSLSSDDL-----DLCLSSANSLY 489
Db 380 AGKSCGAGSVWQAPPEPPDPQVFRMPFDICASHQRSPFF 417

RESULT 8
JC4592
transcription factor ISGF3 gamma chain - mouse
N;Alternate names: interferon-stimulated gene factor 3 gamma chain
C;Species: Mus musculus (house mouse)
C;Date: 10-Apr-1996 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C;Accession: JC4592; S71599
R;Suhara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.
J. Biochem. 119, 231-234, 1996
A;Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma/p48)
A;Reference number: JC4592; MUID:97037063; PMID:8882710
A;Accession: JC4592
A;Molecule type: mRNA
A;Residues: 1-399 <SUH>
A;Cross-references: EMBL:U51992; NID:g1263309; PIDN:AAC52494.1; PID:g1263310
A;Experimental source: L929 cells
R;Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, M.
FEBS Lett. 358, 225-229, 1995
A;Title: Possible involvement of the transcription factor ISGF3gamma in virus-induced exp
A;Reference number: S71599; MUID:95145714; PMID:7843405
A;Accession: S71599
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-87 <KAW>
C;Genetics:
A;Gene: Isgf3g
A;Map position: 14
C;Function:
A;Description: responsible for specific interaction with the promoter element, interferon
ed gene factor 3, the primary regulator of type I interferon responses; involved in type
A;Note: induced by interferon-alpha and interferon-beta
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; signal transduction; transcription factor
F;10-117/Domain: DNA binding #status predicted <DNB>

Query Match 9.5%; Score 259.5; DB 2; Length 399;
Best Local Similarity 22.5%; Pred. No. 9.6e-09;
Matches 105; Conservative 64; Mismatches 192; Indels 105; Gaps 14;

QY 17 WLIGEISSGCEGLQWLDEARTCFRVPWKHFARKDLS-ADARIFKAWAVARGRWPPSSR 75
Db 15 WIVEQVESGHFPQVCMDDAAKTMFRIPWKHAGQDFREDQDAALFKAWALFKK----- 68
QY 76 GGGPPPEAETABAGHKTFRFCALRSTRFVMLRDNSSG-DPADPHKYV-----ALSREL 128
Db 69 -----HKDGDIGHPAVWKTRLCALNKSSEFEEVPERGRMDVABEYKYVYRILPAGTLNPQ 124
QY 129 CWREGPQDTEAEAPAAVPPPGGPPGLFATHAGLQAPGLPAPAGDKGLLQAVQ 188
Db 125 RNQKSPCKRSISCVSPEREBENMNGRTNGVYVNSDSGSGNITGG-----GGNGS----- 171
QY 189 QSCILADHLLTASGADVPVTKAPGEGQGLPLTCACAGGGLPAGELYGWAVETTPSPGP 248
Db 172 -----NRSDSNCSNCELEEG-----AG----- 189
QY 249 QPAALTTEAAAPESPHQAEPYLSPSPSACTAVQEPSPGALDVTIMYKGRVTLQKVVGHP 308
```

```
Db 190 -----TTEATIREDPVFLEHQLPLN-----SDYSLLLTFYGGVVGVKQVHSL 233
Qy 309 SCTFLYGGPDPAVRATDPOQVAPPSPAEPLDQKQRYTEELLRHVAPGLHLRGPQLWA 368
Db 234 DCLVAERSD---SSSMEOVEFPK-----DPLEPTQHLLNQLDRGVLVASNSRGLEFV 284
Qy 369 RMGCKKVYVEGGPPGSPASPACLLPRNCDTPIFDRVFQELVEFRARQRRG-SPR 427
Db 285 QRLCPIPISNWNAPEAPPGGPH-----LLPSNKCVELFKTYFCRDIAQY--FQGGPPPK 338
Qy 428 YTIYLCFGQDLNAGRPKESLVLVKLEPWLCHVHLEGTQREGVSSL 473
Db 339 FOATLHFWEESPSSHSQENLTVQMEQAFARHLLLEKIPERKAAL 384

RESULT 9
S57836
Lymphoid-specific interferon regulator factor - mouse (fragments)
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S57836
R:Matsumura, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Kiefer, P.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A:Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A:Reference number: S57836; MUID:95334364; PMID:7541907
A:Accession: S57836
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-203 <MAT>
A:Cross-references: EMBL:U11692
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 8.2%; Score 224; DB 2; Length 203;
Best Local Similarity 34.8%; Pred. No. 6.7e-07;
Matches 57; Conservative 22; Mismatches 49; Indels 36; Gaps 6;

Qy 16 EWLGEISGCGYEGQLWLDDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74
Db 4 QWLIDQSGKYPGLWENEEKSVFPIPKHAGKQDYNEEDAALFKAWLFGKFP----59
Qy 75 RGGGPPPEAETAEACGKTNFRCALSTRFRFVWLNRNSG-DPADPHKYVALSRELWCWREG 133
Db 60 REGIDKPDPT-----WKTLRLCALNKSNDFFELVERSOLDISDPYKYRI-----105
Qy 134 PGTDQTEAEAPAAVPPGPGPFLATHAGL---QAPGPLPA 174
Db 106 -----VPGQRKNIEKLLSHLERGLVLMWAPDGLYA 136

RESULT 10
T50568
Probable multi-domain regulatory protein [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50568
R:Redenbach, M.; Kiese, H.M.; Denapaita, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: T50568
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1334 <RDS>
A:Cross-references: EMBL:AL133220; PIDN:CA861705.1
A:Experimental source: strain A3(2)
C:Genetics:
A>Note: SCC75A.05c

Query Match 6.7%; Score 183; DB 2; Length 1334;
Best Local Similarity 28.1%; Pred. No. 0.0015;
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;
```

```
Qy 18 LLGEISSGCGYEGQLWLDDEARTCFRVPWKHFARKDLSADARIFKAWAVARGRWPPSSRG 77
Db 146 VLADLPDRTAEAAR--ETRHFEALRARHTAALDQGAHSLPELTALCDGH-----195
Qy 78 GPPPEAETAEER-----AGWKNFRCALSTRFRFVWLNRNSG-DPA-----D 117
Db 196 -PLDEPLQALRLRALSRTAEALAAVEARR--LLADRLGTDGPELRTLHAELLSPS 252
Qy 118 PHKYVALSRELWCWREG-----GTDQTEAEAPAAVPPPPQGGP-PGPFALATHAGL 166
Db 253 PIPTFGSRSTPGWTSPPGSPASGAGAASGTDVASGAGAAGDPDPASGPASGPAVPGSGGG 312
Qy 167 QAPGELPAPAGDKGDLLOAVQOOSCLADHLLTASWGADVPVTKAPCGEGGGLPTGCACAG 226
Db 313 PAPGWPAPAGTAPGSGSTAPPHDTASAAD---TA---PAPGPTSAPGTA---PAGTAAP 362
Qy 227 GPGI--PA-GELY--GWA--VETTPSPGPQPAALTTGEA-----AAPESPHQAEPLY 271
Db 363 APGTAGPAGTGYAPGTAPVAGTTTAPGTAPAGTAPAGTAPAGTAPAGTAPAGTAPAGT 422
Qy 272 SPSPSACTAV-----QEPSPGALDVTIMYKGRITVLQKVGHPSCTFLYGPDPVAVRAT 324
Db 423 APAPGSTPAGTVPAPGTAPAPGQPA---DGR---RPVTGPASCT---GPG---AAT 468
Qy 325 DPQOVA-----FPSPA 335
Db 469 PPEAAAAAASAGSAPSPA 485

RESULT 11
I52998
Interferon regulatory factor 1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I52998
R:Cha, Y.; Sims, S.H.; Romine, M.F.; Kaufmann, M.; Deiseroth, A.B.
DNA Cell Biol. 11, 605-611, 1992
A:Title: Human interferon regulatory factor 1: intron/exon organization.
A:Reference number: I52998; MUID:93000481; PMID:1382447
A:Accession: I52998
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: GB:L05072; NID:g184648; PIDN:AAA36043.1; PID:g184649
C:Genetics:
A:Genes: GDB:IRF1
A:Cross-references: GDB:127269; OMIM:147575
A:Map position: 5q31.1-5q31.1
A:Introns: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1

Query Match 6.4%; Score 176; DB 2; Length 325;
Best Local Similarity 23.6%; Pred. No. 0.0089;
Matches 69; Conservative 35; Mismatches 108; Indels 80; Gaps 12;

Qy 11 RVLFGELWLGTEISSGCGYEGQLWLDDEARTCFRVPWKHFARK--DLSEADARIFKAWAVARG 68
Db 5 RMRMRPWLEMQINSQIPLGLIWINKEEMIFQIPWKHAAKHGWDINK-DACLFERSWAIHTG 63
Qy 69 RNPSPSRGCGPPPEAETAEACGKTNFRCALSTRFRFVWLNRD---NSGDPADPHKYVALS 125
Db 64 RY-----KAGEKPDPT-----WKANFRCAANSLDIEEVKQOSNRKNGSSA--VRVYRML 112
Qy 126 RELC---WREGPGTDQTEAEAPAAVPPPGPGPF-----LAHTHAGLQAPGPLPA 174
Db 113 PFLTNNQKERSKSRDAKSKAKRS CGDSSPDIFSDGLSSSTLPDDHSSYTVPG---168
Qy 175 PAGDKGDLLOAVQOOSCLADHLLTASWGADVPVTKAPGEGQEGELPTGCACAGGGLP---231
Db 169 -----YMQDLEVEQALT-----PALSPCAVSSTLPDWH 196
Qy 232 -----AGELYGWAVETTPSPGPQPAALTTGEAAAPESP HQAEPLYSPS 274
Db 197 IPVEVVPDSTSDLYNFQV-----SPMPSTSEATTDEDEGKLPEIDIMKLEQS 244
```

RESULT 12
A39587
collagen, cornea-specific - chicken
C;Species: Gallus gallus (chicken)
C;Date: 14-Feb-1992 #sequence_revision 15-Aug-1997 #text_change 20-Sep-1999
C;Accession: S16501; A38587
R;Marchant, J.K.; Linseman, T.F.; Gordon, M.K.
Proc. Natl. Acad. Sci. U.S.A. 88, 1560-1564, 1991
A;Title: cDNA analysis predicts a cornea-specific collagen.
A;Reference number: A38587; MUID:91142213; PMID:1705041
A;Accession: S16501
A;Molecule type: mRNA
A;Residues: 1-1146 <MAR>
A;Cross-references: EMBL:M60172; NID:G211609; PIDN:AAA48703.1; PID:G211610
A;Accession: A38587
A;Molecule type: mRNA
A;Residues: 1-174, 'X', 176-233, 'X', 235-344, 'X', 346-408, 'X', 410-499, 'X', 501-876, 'X', 878-1111
A;Cross-references: GB:M60172
C;Superfamily: unassigned collagens
C;Keywords: cornea

Query Match 6.4%; Score 176; DB 2; Length 1146;
Best Local Similarity 23.8%; Pred. No. 0.0033;
Matches 109; Conservative 31; Mismatches 132; Indels 186; Gaps 26;

QY 66 ARGRWPPSSRGG-----GPPPEAETAEARAGKTNFRFCALRSTRFRVFLNRDSDGPAD 117
Db 348 ARGPFGSGDTGTFGTGPGQPGPLPQNGRPGAK-----GEPGA 387

QY 118 PKVYALSRELWCWREGPGTDTAEAPAAVPPQGGPGPLAHTAGLQAP-GP--LP- 173
Db 388 PKVIS-----AGSSTIALPGPP-GPPGIGTGPVPGVPGAGLPG 431

QY 174 --APAGDKG---DLLLQAV--QOCLADHLLTASWG--ADPVPTKAPGEGQGLPLTGAC 224
Db 432 QQGPGEKGSAAVEWETIKTEVSSLASQMLSDQGRAGPPGPPGESVQGLP---GP 488

QY 225 AGGFLPAGELYGWAVETPS--PGQPAAITGEE--AAAPESP-----HQAPYLS 272
Db 489 RGPPLPGF-----PSGPPGRSGSVSTSETFVSGPPGPPGPKGQGE----- 533

QY 273 PPSACTAVOEPSFGALDVTIMKGRVTLQKVGHFSCFELYGPPDPVAVRATPP----- 326
Db 534 PGPRTGTG--EPGEPGLPGSSHGVTVMQPGPP-----GPPGKGDAGVPGAPGIP 585

QY 327 -----QQV-----AFPSPAELPDQQLRYTEELL-----RHVAPGLHLELRGP 364
Db 586 GTSRGGSRQIQGPPGPPGPPGPPGPGGSSQEIQQYVADYLSKSDNVRHYLTG----- 637

QY 365 QLWARMGKCKVYVEVCGPPGPGSASPSTPACLLPNCPTPIFDF----- 407
Db 638 -----VQGPFG--PGPPGIL--TTADGKNFDFAEALATRVMSYVTSDDH 678

QY 408 -----RVFQBLVEFRARQR-----RGSP 426
Db 679 YQSFASVSTSVLYQELNMLQREIRQLVGRGPP 716

RESULT 13
A53340
interferon regulatory factor 2 - human
N;Alternate names: transcription repressor IRF2
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
C;Accession: A53340; S06894; A32828
R;Cha, Y.; Deisseroth, A.B.
J. Biol. Chem. 269, 5279-5287, 1994
A;Title: Human interferon regulatory factor 2 gene. Intron-exon organization and function.
A;Reference number: A53340; MUID:94148994; PMID:8106512
A;Accession: A53340
A;Status: not compared with conceptual translation

A;Molecule type: DNA
A;Residues: 1-349 <CHA>
A;Cross-references: GB:L24442
R;Itoh, S.; Harada, H.; Fujita, T.; Mimura, T.; Taniguchi, T.
Nucleic Acids Res. 17, 8372, 1989
A;Title: Sequence of a cDNA coding for human IRF-2.
A;Reference number: S06894; MUID:90045964; PMID:2813069
A;Accession: S06894
A;Molecule type: mRNA
A;Residues: 1-57, 'R', 59-349 <ITO>
A;Cross-references: EMBL:X15949; NID:G33966; PIDN:CAA34073.1; PID:G33967
R;Harada, H.; Fujita, T.; Miyamoto, M.; Kimura, Y.; Maruyama, M.; Furia, A.; Miyata, T.; Cell 58, 729-739, 1989
A;Title: Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to DNA.
A;Reference number: A32828; MUID:89354547; PMID:2475256
A;Accession: A32828
A;Molecule type: mRNA
A;Residues: 1-68, 'I', 70-96, 'R', 98-130, 'ER', 133-152, 'GF', 155-163, 'A', 165-188, 'D', 190-210, 'R', 210-313, 'PAPV', 318-319, 'TP', 322-349 <HAR>
A;Cross-references: GB:J03168; NID:G198455; PIDN:AAA39333.1; PID:G293676
A;Experimental source: clones 2 and 5
C;Genetics:
A;Gene: GDB:IRF2
A;Cross-references: GDB:127270; OMIM:147576
A;Map position: 4q35.1-4q35.1
C;Keywords: DNA binding; transcription regulation

Query Match 6.1%; Score 167; DB 2; Length 349;
Best Local Similarity 35.0%; Pred. No. 0.0034;
Matches 36; Conservative 18; Mismatches 39; Indels 10; Gaps 3;

QY 11 RVLFGWLLGESSGCGEGLQWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGR 69
Db 5 RRMRRPWLREQINSNTIPGLKWLNEKKIFQIPWMAAHHGWDEKDALFRWVAHTGK 64

QY 70 WPPSSRGGGPPPEAETAEARAGKTNFRFCALRSTRFRVFLNRDLS 112
Db 65 HQP-----GVDPKDPKT-----WKNFRCAVNSLPDIEEVKDKS 98

RESULT 14
B31595
interferon regulatory factor 1 - human
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 16-Feb-1997
C;Accession: B31595; S04075
R;Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.; Cell 54, 903-913, 1988
A;Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifies cell fate.
A;Reference number: A90903; MUID:88311092; PMID:3409321
A;Accession: B31595
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-325 <M12>
R;Maruyama, M.; Fujita, T.; Taniguchi, T.
Nucleic Acids Res. 17, 3292, 1989
A;Title: Sequence of a cDNA coding for human IRF-1.
A;Reference number: S04075; MUID:89263736; PMID:2726461
A;Accession: S04075
A;Molecule type: mRNA
A;Residues: 1-325 <MAR>
A;Cross-references: EMBL:X14454
C;Keywords: DNA binding; transcription regulation

Query Match 6.1%; Score 166; DB 2; Length 325;
Best Local Similarity 23.4%; Pred. No. 0.0036;
Matches 67; Conservative 33; Mismatches 106; Indels 80; Gaps 12;

QY 17 WLIGEISSGCGEGLQWLDEARTCFRVPWKHFARK--DLSEADARIFKAWAVARGRWPSS 74
Db 11 WLEMQINSQIPLGLIWNKEEMILEIPWMAAHHGWDEKDALFRWVAHTGRI----- 65

QY 75 RGGGPPPEAETAEARAGKTNFRFCALRSTRFRVFLNRDLS---NSGDPADPHKYALSRELCL-- 129


```
Db      66 KAGEKEPDPT-----WKNFRCAAMSLPDIIEVKDQSRNKGSSA--VRVYRMLPPLTKN 118
Qy      130 -WREGFGTDQTEAEAPAAVPPPGGPPGF-----LAHTHAGLQAPGPLPAPAGDKG 180
Db      119 QRKERKSSRDASKAKRKSCKDSSPDFTDGLSSSTLPDDHSSYTVFG-----168
Qy      181 DILLQAVQSCSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLP-----231
Db      169 -----YMODLEVEQALT-----PALSPCAVSVSTLPDWHIPVEVV 202
Qy      232 ---AGELYGMVAVETTPSPGPQPAALTGEEAAAPESPHQAEPVLSPS 274
Db      203 PDSTSLYNFQV-----SPMPSISEATTDEEGKLPEDIMKLEQS 244

RESULT 15
A31595
interferon regulatory factor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 05-Nov-1999
C:Accession: A31595
R:Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;
Cell 54, 903-913, 1988
A:Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifica
A:Reference number: A90903; MUID:88311092; PMID:3409321
A:Accession: A31595
A:Molecule type: mRNA
A:Residues: 1-329 <MIY>
A:Cross-references: GB:M21065; NID:gl98458; PIDN:AAA39334.1; PID:9293677; GB:J03160; GB:
C:Keywords: DNA binding; transcription regulation
```

```
Query Match      6.1%; Score 166; DB 2; Length 329;
Best Local Similarity 21.6%; Pred. No. 0.0036;
Matches 65; Conservative 40; Mismatches 88; Indels 108; Gaps 12;

Qy      11 RVLFGEWLLGEISSGCGYEGLOWLDEARTCFRVPWKHFARK--DLSEADARIFKAWAVARG 68
Db      5  RMRMRPWLEMQINSQIPLGLIWINKEEMIFOIPWKHAAKHGWDINK-DACLFRSWAIHTG 63
Qy      69 RWPSSRGGPPPEAETAEAGWKTNFRCALRSTRFVMLRD--NSGDPADPHKVYAL- 124
Db      64 RY-----KAGEKEPDPT-----WKNFRCAAMSLPDIIEVKDQSRNKGSSA--VRVYRML 112
Qy      125 -----SRELWREGFGT-----136
Db      113 PPLTRNQRKERKSSRDTKSKTKKLCGVSPDFTSDGLSSSTLPDDHSSYTTQGYLQ 172
Qy      137 -----DQTEAEAPAAVPPPGGPPGFFLAHTHAGLQ-----APGPLPAPA 176
Db      173 DLDMERDITPALLSPCVV-----SSSLSEWHMQMDIIPDSTTDLYNLQVSPMPSTSEA 224
Qy      177 GDKGDLQLQAVQSCSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELY 236
Db      225 ATDED-----BEGKTAEDIMKLFEQSEWQPTHIDGKY-----LLNEFGTQLSSVY 270
Qy      237 G 237
Db      271 G 271
```

Search completed: June 18, 2003, 12:48:17
Job time : 24 secs

THIS PAGE BLANK (USPTO)

CC and threonine residues in the carboxy terminus are modified post-
CC translationally, by phosphorylation, following sendai virus infection.
CC The modified IRF-7 substituted with aspartic acid, functions as a
CC activator of promoters containing ISRE regulatory elements and
CC stimulation of DNA binding and transcriptional activity. IRF-7 protein
CC shows highest homology to IRF-3. The modified IRF proteins are used for
CC the treatment of viral infections like, influenza, herpes or HIV
CC infection. They may also be used to activate a cytokine gene, in cancer
CC treatment or to modify a target cell of an organism
XX
SQ Sequence 503 AA;

Query Match 100.0%; Score 2731; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 1-2e-196;
Matches 503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALAPERAPRVLFGEWLLGEISGGCYEGQLWDEARTCFRVPWKHPARKDLSEADARIF 60
DB 1 MALAPERAPRVLFGEWLLGEISGGCYEGQLWDEARTCFRVPWKHPARKDLSEADARIF 60

QY 61 KAWAVARGRWPSPSRGGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
DB 61 KAWAVARGRWPSPSRGGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120

QY 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGFLPAPADK 180
DB 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGFLPAPADK 180

QY 181 DILLQAVQOQSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DILLQAVQOQSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240

QY 241 ETTSPGQPOPAALTGTEAAAPESP HQAEPLYLSPSPSACTAVQSPGALDVTIMYKGRV 300
DB 241 ETTSPGQPOPAALTGTEAAAPESP HQAEPLYLSPSPSACTAVQSPGALDVTIMYKGRV 300

QY 301 LQKVVGHPSCFTLYGPPDPDAVRATDPOQVAFPPGPAELPDQKQRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTLYGPPDPDAVRATDPOQVAFPPGPAELPDQKQRYTEELLRHVAPGLHLE 360

QY 361 LRGPQLWARRMGKCKVYVEGPPGSGASPSSTPACLLPRNCDTPIFDPRVFQELVEPRAR 420
DB 361 LRGPQLWARRMGKCKVYVEGPPGSGASPSSTPACLLPRNCDTPIFDPRVFQELVEPRAR 420

QY 421 QRRGSPRYTYLFGQDLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSLDL 480
DB 421 QRRGSPRYTYLFGQDLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSLDL 480

QY 481 CLSSANSLYDDIECFMLEQPA 503
DB 481 CLSSANSLYDDIECFMLEQPA 503

RESULT 2

ADJ75462
ID ADJ75462 standard; protein; 503 AA.

XX AC ADJ75462;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene related amino acid sequence SEQ ID NO:714.

XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX KW gene therapy; marker.

XX OS Homo sapiens.

XX PN EPI394274-A2.

XX PD 03-MAR-2004.

XX

PF 04-AUG-2003; 2003EP-00254857.
XX
PR 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
PA (GENO-) GENOX RES INC.
XX
PI Ontani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
DR

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 714; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC invention is used in the exemplification of the present invention.

XX SQ Sequence 503 AA;

Query Match 99.6%; Score 2719; DB 8; Length 503;
Best Local Similarity 99.6%; Pred. No. 9.5e-196;
Matches 501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALAPERAPRVLFGEWLLGEISGGCYEGQLWDEARTCFRVPWKHPARKDLSEADARIF 60

DB 1 MALAPERAPRVLFGEWLLGEISGGCYEGQLWDEARTCFRVPWKHPARKDLSEADARIF 60

QY 61 KAWAVARGRWPSPSRGGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120

DB 61 KAWAVARGRWPSPSRGGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120

QY 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGFLPAPADK 180

DB 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGFLPAPADK 180

QY 181 DILLQAVQOQSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240

DB 181 DILLQAVQOQSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240

QY 241 ETTSPGQPOPAALTGTEAAAPESP HQAEPLYLSPSPSACTAVQSPGALDVTIMYKGRV 300

DB 241 ETTSPGQPOPAALTGTEAAAPESP HQAEPLYLSPSPSACTAVQSPGALDVTIMYKGRV 300

QY 301 LOKVGHPSCTFLYGPDPVAVRATDQQVAFSPAEPLDPQKQRYTEELLRHVAPGLHLE 360
DB 301 LOKVGHPSCTFLYGPDPVAVRATDQQVAFSPAEPLDPQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPOLWARRMGKCKVYWEVGGPGSASPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
DB 361 LRGPOLWARRMGKCKVYWEVGGPGSASPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLVLEPWLKRVHLEGTQREGVSSLDSSLDL 480
DB 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLVLEPWLKRVHLEGTQREGVSSLDSSLDL 480
QY 481 CLSSANSLYDDIECFMELEQPA 503
DB 481 CLSSANSLYDDIECFMELEQPA 503
RESULT 3
ADJ75410
ID ADJ75410 standard; protein; 503 AA.
AC ADJ75410;
XX
XX
XX 20-MAY-2004 (first entry)
XX
DE Marker gene related amino acid sequence SEQ ID NO:662.
XX
XX bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.
XX
XX Homo sapiens.
XX
XX EP1394274-A2.
XX
XX 03-MAR-2004.
XX
XX 04-AUG-2003; 2003EP-00254857.
XX
XX 06-AUG-2002; 2002JP-00229312.
XX
XX 20-MAR-2003; 2003JP-00077212.
XX
XX (GENO-) GENOX RES INC.
XX
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX
XX WPI; 2004-193155/19.
XX
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX comparing the expression level of a marker gene in a biological sample
XX from a subject with the expression level of the gene in a sample from a
XX healthy subject.
XX
XX Example 11; SEQ ID NO 662; 241pp; English.
XX
XX The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic

CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 503 AA;

Query Match 99.6%; Score 2719; DB 8; Length 503;
Best Local Similarity 99.6%; Pred. No. 9.5e-196;
Matches 501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALAPERAAPRVLFGEWLLGSEISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGSEISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGGGPPPEAETARAGWKTNFRCLARSTRFVMLRDNDSGDPADPHK 120
DB 61 KAWAVARGWPPSSRGGGPPPEAETARAGWKTNFRCLARSTRFVMLRDNDSGDPADPHK 120
QY 121 VVALSRELWCWREGPCTDQTEAEPAAVPPQCGGPPGFLAHTAGLQAPGFLPAPAGDKG 180
DB 121 VVALSRELWCWREGPCTDQTEAEPAAVPPQCGGPPGFLAHTAGLQAPGFLPAPAGDKG 180
QY 181 DILLQAVQOCSLADHLLTASWGADVPVTKAPGEGQEGPLTGCACAGGPGLPAGELYGWAV 240
DB 181 DILLQAVQOCSLADHLLTASWGADVPVTKAPGEGQEGPLTGCACAGGPGLPAGELYGWAV 240
QY 241 ETTSPGPGPQALTTGEAAAPESPHOAEPLYSPSPSACTAVOEPSGALDVTIMYKGRV 300
DB 241 ETTSPGPGPQALTTGEAAAPESPHOAEPLYSPSPSACTAVOEPSGALDVTIMYKGRV 300
QY 301 LOKVGHPSCTFLYGPDPVAVRATDQQVAFSPAEPLDPQKQRYTEELLRHVAPGLHLE 360
DB 301 LOKVGHPSCTFLYGPDPVAVRATDQQVAFSPAEPLDPQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPOLWARRMGKCKVYWEVGGPGSASPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
DB 361 LRGPOLWARRMGKCKVYWEVGGPGSASPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLVLEPWLKRVHLEGTQREGVSSLDSSLDL 480
DB 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLVLEPWLKRVHLEGTQREGVSSLDSSLDL 480
QY 481 CLSSANSLYDDIECFMELEQPA 503
DB 481 CLSSANSLYDDIECFMELEQPA 503

RESULT 4
ADE58665
ID ADE58665 standard; protein; 503 AA.
XX
AC ADE58665;
XX
DT 29-JAN-2004 (first entry)
XX
XX Human Protein Q92985, SEQ ID NO 4541.
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
PN

```
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q92985.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a human protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 503 AA;
Query Match 99.4%; Score 2715; DB 7; Length 503;
Best Local Similarity 99.4%; Pred. No. 1.9e-195;
Matches 500; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALAPERAAFRVLFGWLLGEISSCGVEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAFRVLFGWLLGEISSCGVEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAMAVARGRPPSSRRGGPPPEAFTAEARAGKTNFRCAKSTRFRFVMLRDSNGDPADPHK 120
DB 61 KAMAVARGRPPSSRRGGPPPEAFTAEARAGKTNFRCAKSTRFRFVMLRDSNGDPADPHK 120
QY 121 VYALSRLCWREGSGTQTEAEAPAAVPPQGGPPGPFLLATHAGLQAPGFLPAPADKG 180
DB 121 VYALSRLCWREGSGTQTEAEAPAAVPPQGGPPGPFLLATHAGLQAPGFLPAPADKG 180
QY 181 DLLLQAVQOSCLADHLLTASNGADPVPTKAPGEGQGLPTGACAGGPGLPAGELYGNAV 240
DB 181 DLLLQAVQOSCLADHLLTASNGADPVPTKAPGEGQGLPTGACAGGPGLPAGELYGNAV 240
QY 241 ETTTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSPSACTAVQSPSGALDVTIMYKGRTV 300
DB 241 ETTTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSPSACTAVQSPSGALDVTIMYKGRTV 300
QY 301 LQKVVGHPSCFTLYGPPDPAVRATDPQVAFPSPAELPQOKQLRYTEELRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTLYGPPDPAVRATDPQVAFPSPAELPQOKQLRYTEELRHVAPGLHLE 360
QY 361 LRGPOLWARRMGKCKVYWEVGGPPGSASPSTPACLLPRNCDTPIFDFRVFFQELVEFRAR 420
DB 361 LRGPOLWARRMGKCKVYWEVGGPPGSASPSTPACLLPRNCDTPIFDFRVFFQELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLGAGRPKEKSLVLKLEPWLRCVHLEGTQREGVSSLDSSDL 480
DB 421 QRRGSPRYTIYLGFGQDLGAGRPKEKSLVLKLEPWLRCVHLEGTQREGVSSLDSSDL 480
QY 481 CUSSANSLYDDIECFLMELEQPA 503
DB 481 CUSSANSLYDDIECFLMELEQPA 503
RESULT 5
AD58669
ID AD58669 standard; protein; 503 AA.
XX AC ADE58669;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein Q92985, SEQ ID NO 4545.
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WC2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q92985.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a human protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
```


RESULT 7
ADJ75465
ID ADJ75465 standard; protein; 516 AA.
XX AC ADJ75465;
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene related amino acid sequence SEQ ID NO:717.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX KW gene therapy; marker.
XX OS Homo sapiens.
XX PN EP1394274-A2.
XX PD 03-MAR-2004.
XX PF 04-AUG-2003; 2003EP-00254857.
XX PR 06-AUG-2002; 2002JP-00229312.
XX PR 20-MAR-2003; 2003JP-00077212.
XX PA (GENO-) GENOX RES INC.
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX PT comparing the expression level of a marker gene in a biological sample
XX PT from a subject with the expression level of the gene in a sample from a
XX PT healthy subject.
XX PS Example 11; SEQ ID NO 717; 241pp; English.
XX CC The present invention describes a method of testing for bronchial asthma
XX CC or chronic obstructive pulmonary disease. The method comprises
XX CC determining the expression level of a marker gene in a biological sample
XX CC from a subject, comparing the expression level determined with the
XX CC expression level of the marker gene in a biological sample from a healthy
XX CC subject, and judging whether the subject has bronchial asthma or chronic
XX CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX CC genes (S1) whose expression levels increase when respiratory epithelial
XX CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX CC whose expression levels decrease when respiratory epithelial cells are
XX CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX CC testing for bronchial asthma or chronic obstructive pulmonary disease;
XX CC (2) a kit for screening for a candidate compound for a therapeutic agent
XX CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX CC an animal model for bronchial asthma or chronic obstructive pulmonary
XX CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX CC method for producing an animal model for bronchial asthma or chronic
XX CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX CC asthma or chronic obstructive pulmonary disease, comprising the compound,
XX CC a marker gene or an antisense nucleic acid corresponding to a portion of
XX CC the marker gene, a ribozyme, a polynucleotide that suppresses the
XX CC expression of the gene through an RNAi effect or an antibody recognising
XX CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX CC probe has been immobilised to assay a marker gene. (I) has respiratory
XX CC and antiasthmatic activities, and can be used in gene therapy. The method
XX CC is useful for testing for or screening for a therapeutic agent for
XX CC bronchial asthma or chronic obstructive pulmonary disease. The present
XX CC sequence is used in the exemplification of the present invention.
XX SQ Sequence 516 AA;

Query Match 98.2%; Score 2682; DB 8; Length 516;
Best Local Similarity 99.2%; Pred. No. 5.9e-193;

	Matches	493;	Conservative	2;	Mismatches	2;	Indels	0;	Gaps	0;
QY	7	RAAPRVLFGEWLLGEIS	SSCYEGLOWLDEARTCF	RVPMKHFARKDI	SEADARIFKAWA	66				
DB	20	RAAPRVLFGEWLLGEIS	SSCYEGLOWLDEARTCF	RVPMKHFARKDI	SEADARIFKAWA	79				
QY	67	RGRWPPSSRGGGPPPEA	ETAEAGWKTNFR	CFALRSTRFV	MLRDN	SGDPADPHKVALSR	126			
DB	80	RGRWPPSSRGGGPPPEA	ETAEAGWKTNFR	CFALRSTRFV	MLRDN	SGDPADPHKVALSR	139			
QY	127	ELCWREGPGTDQTEA	BAPAAVPPPGG	PPFLAHTAGL	QAPGPI	PAPAGDKDLLLQA	186			
DB	140	ELCWREGPGTDQTEA	BAPAAVPPPGG	PPFLAHTAGL	QAPGPI	PAPAGDKDLLLQA	199			
QY	187	VQOSCLADHLLTAS	WGADPVTTPA	PGEGQGLPLT	GACAGGPG	LPA	246			
DB	200	VQOSCLADHLLTAS	WGADPVTTPA	PGEGQGLPLT	GACAGGPG	LPA	259			
QY	247	GPOPAALTTGEAA	AAAPESPHQAE	PYLSPSACTA	VOEPS	PGALDVTIMYKGR	306			
DB	260	GPOPAALTTGEAA	AAAPESPHQAE	PYLSPSACTA	VOEPS	PGALDVTIMYKGR	319			
QY	307	HPSCFTLYGPPDPA	VRATDQQVAF	PSPAELPQK	OLRYTEEL	LRHVAPGLHLELR	366			
DB	320	HPSCFTLYGPPDPA	VRATDQQVAF	PSPAELPQK	OLRYTEEL	LRHVAPGLHLELR	379			
QY	367	WARMGKCKYWEV	GGPGSGAS	SPSTACLL	PRNC	DTPI	426			
DB	380	WARMGKCKYWEV	GGPGSGAS	SPSTACLL	PRNC	DTPI	439			
QY	427	RYTYILGFGQDLS	AGRPKEKSLV	LKLEPWL	CRVHLEGT	QREGVSSLD	486			
DB	440	RYTYILGFGQDLS	AGRPKEKSLV	LKLEPWL	CRVHLEGT	QREGVSSLD	499			
QY	487	SLYDDIECFLE	MELEQPA	503						
DB	500	SLYDDIECFLE	MELEQPA	516						

RESULT 8
ADJ75413
ID ADJ75413 standard; protein; 516 AA.
XX AC ADJ75413;
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene related amino acid sequence SEQ ID NO:665.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX KW gene therapy; marker.
XX OS Homo sapiens.
XX PN EP1394274-A2.
XX PD 03-MAR-2004.
XX PF 04-AUG-2003; 2003EP-00254857.
XX PR 06-AUG-2002; 2002JP-00229312.
XX PR 20-MAR-2003; 2003JP-00077212.
XX PA (GENO-) GENOX RES INC.
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX PT comparing the expression level of a marker gene in a biological sample
XX PT from a subject with the expression level of the gene in a sample from a

PT healthy subject.

PS Example 11; SEQ ID NO 665; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (I) has respiratory
XX and antiasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.

SQ Sequence 516 AA;

Query Match 98.2%; Score 2682; DB 8; Length 516;

Best Local Similarity 99.2%; Pred. No. 5.9e-193;

Matches 493; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVLFGEWLLGEISSGCEYGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVA 66
DB 20 RAAPRVLFGEWLLGEISSGCEYGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVA 79
QY 67 RGRWPPSSRGSGGPPPPAEATAERAGWKNFRCALRSTRFRVWMLRDNSGDDPADPHKVYALSR 126
DB 80 RGRWPPSSRGSGGPPPPAEATAERAGWKNFRCALRSTRFRVWMLRDNSGDDPADPHKVYALSR 139
QY 127 ELCWREGPGTDTAEAPAAVPPQGGPPGFPLAHTHAGLQAPGPLPAPAGDKGDLLOA 186
DB 140 ELCWREGPGTDTAEAPAAVPPQGGPPGFPLAHTHAGLQAPGPLPAPAGDEGDLLOA 199
QY 187 VQOSCLADHLLTASWADPVTTPAPGEGQGLTTCACAGGPGLPAGELYGNAVETTPSP 246
DB 200 VQOSCLADHLLTASWADPVTTPAPGEGQGLTTCACAGGPGLPAGELYGNAVETTPSP 259
QY 247 GPQPAALTGTEAAAPSPHQAEVPLSPSACTAVQSPSGALDVTIMYKGRVLOKVVG 306
DB 260 GPQPAALTGTEAAAPSPHQAEVPLSPSACTAVQSPSGALDVTIMYKGRVLOKVVG 319
QY 307 HPSCFTLYGPPDPVAVRATPDQVAFPSPAELPDQKQLRYTELLRHVAFGLHLELPGQL 366
DB 320 HPSCFTLYGPPDPVAVRATPDQVAFPSPAELPDQKQLRYTELLRHVAFGLHLELPGQL 379
QY 367 WARMGCKVYVWVGPPGASSTPACLLPRNCDTPIFDPRVFFQELVEFRARQRGSP 426
DB 380 WARMGCKVYVWVGPPGASSTPACLLPRNCDTPIFDPRVFFQELVEFRARQRGSP 439
QY 427 RYTIYLGFGODLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDLCLSSAN 486
DB 440 RYTIYLGFGODLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDLCLSSAN 499
QY 487 SLYDDIECFLEMELEQPA 503

DB 500 SLYDDIECFLEMELEQPA 516

RESULT 9

ADL83070

ID ADL83070 standard; protein; 516 AA.

XX ADL83070;

DT 17-JUN-2004 (first entry)

DE Human PRO83673, SEQ ID 272.

XX Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic;
KW Antiallergic; Muscular; Neuroprotective; Nephrotropic; Antiinflammatory;
KW Gene Therapy; PRO; B cell related disorder; cancer;
KW immune-mediated inflammatory disease; human.

OS Homo sapiens.

PN WO2004024097-A2.

PD 25-MAR-2004.

PF 15-SEP-2003; 2003WO-US029097.

PR 16-SEP-2002; 2002US-0411392P.

XX (GETH) GENENTECH INC.

PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood W;

PI Wu TD;

XX WPI: 2004-329389/30.

DR N-PSDB; ADL83069.

XX New PRO polypeptide, useful for diagnosing and treating a B cell related
PT disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune
PT mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.

PS Claim 10; Fig 272; 695pp; English.

XX The present invention relates to PRO proteins and their coding sequences.
CC The PRO proteins are useful for diagnosing and treating a B cell related
CC disorder, e.g. X-linked infantile hypogammaglobulinemia, polyaccharide
CC antigen unresponsiveness, selective IgA deficiency, selective IgM
CC deficiency, selective deficiency of IgG subclasses, immunodeficiency with
CC hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's
CC lymphoma, intermediate lymphoma, follicular lymphoma, type II
CC hypersensitivity, rheumatoid arthritis, autoimmune mediated haemolytic
CC anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or
CC ankylosing spondylitis. The PRO proteins are also useful for preparing a
CC medicament for treating a condition that is responsive to the PRO
CC protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO
CC coding sequences are useful as hybridization probes in chromosome and
CC gene mapping, in preparing PRO proteins, or in generating transgenic
CC animals or knockout animals, which in turn are useful in the development
CC and screening of therapeutically useful reagents.

XX Sequence 516 AA;

Query Match 98.2%; Score 2682; DB 8; Length 516;

Best Local Similarity 99.2%; Pred. No. 5.9e-193;

Matches 493; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVLFGEWLLGEISSGCEYGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVA 66

DB 20 RAAPRVLFGEWLLGEISSGCEYGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVA 79

QY 67 RGRWPPSSRGSGGPPPPAEATAERAGWKNFRCALRSTRFRVWMLRDNSGDDPADPHKVYALSR 126

DB 80 RGRWPPSSRGSGGPPPPAEATAERAGWKNFRCALRSTRFRVWMLRDNSGDDPADPHKVYALSR 139

QY 127 ELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPLPAPAGDKGDLLOA 186
DB 140 ELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPLPAPAGDEGDLLOA 199
QY 187 VQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 246
DB 200 VQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 259
QY 247 GPQPAALTTGEEAAAPSPHOAEPLYLSPSPSACTAVQEPSGALDVTIMYKGRVLOKVVG 306
DB 260 GPQPAALTTGEEAAAPSPHOAEPLYLSPSPSACTAVQEPSGALDVTIMYKGRVLOKVVG 319
QY 307 HPSCCTFLYGPDPFAVRATDPQQVAFPSPAELPDQKQRYTEELRHVAPGLHLELRGPOL 366
DB 320 HPSCCTFLYGPDPFAVRATDPQQVAFPSPAELPDQKQRYTEELRHVAPGLHLELRGPOL 379
QY 367 WARMGKCKYWEVGGPPGSGASSTPACLLPRNCPTPIFDRVFFQELVEFRARORRGP 426
DB 380 WARMGKCKYWEVGGPPGSGASSTPACLLPRNCPTPIFDRVFFQELVEFRARORRGP 439
QY 427 RYTIYLGFGQDLSAGRPKEKSLVLEPWLCRVHLEGTQREGVSSLDSSDLCLSSAN 486
DB 440 RYTIYLGFGQDLSAGRPKEKSLVLEPWLCRVHLEGTQREGVSSLDSSDLCLSSAN 499
QY 487 SLYDDIECFMLELEQPA 503
DB 500 SLYDDIECFMLELEQPA 516
RESULT 10
ADN04327
ID ADN04327 standard; protein; 516 AA.
AC ADN04327;
XX
XX 01-JUL-2004 (first entry)
XX Antipsoriatic protein sequence #358.
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.
XX
XX WO2004028479-A2.
XX
XX 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH) GENENTECH INC.
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WJ;
PI Wu TD;
XX WPI; 2004-305105/28.
XX N-PSDB; ADN04326.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
XX Claim 9; SEQ ID NO 721; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
XX Sequence 516 AA;
SQ

Query Match 98.2%; Score 2682; DB 8; Length 516;
Best Local Similarity 99.2%; Pred. No. 5.9e-193;
Matches 493; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 7 RAAPRVLFGEWLLGEISSGCEGLQWLDEARTCFRVPWKHFARKDISEADARIFKAWAVA 66
DB 20 RAAPRVLFGEWLLGEISSGCEGLQWLDEARTCFRVPWKHFARKDISEADARIFKAWAVA 79
QY 67 RGRWPPSSGGGPPPPAEATAERAGWKTNFRCALRSTRFRVWLNRDNGDDPADPHKVVALSR 126
DB 80 RGRWPPSSGGGPPPPAEATAERAGWKTNFRCALRSTRFRVWLNRDNGDDPADPHKVVALSR 139
QY 127 ELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPLPAPAGDKGDLLOA 186
DB 140 ELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPLPAPAGDEGDLLOA 199
QY 187 VQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 246
DB 200 VQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 259
QY 247 GPQPAALTTGEEAAAPSPHOAEPLYLSPSPSACTAVQEPSGALDVTIMYKGRVLOKVVG 306
DB 260 GPQPAALTTGEEAAAPSPHOAEPLYLSPSPSACTAVQEPSGALDVTIMYKGRVLOKVVG 319
QY 307 HPSCCTFLYGPDPFAVRATDPQQVAFPSPAELPDQKQRYTEELRHVAPGLHLELRGPOL 366
DB 320 HPSCCTFLYGPDPFAVRATDPQQVAFPSPAELPDQKQRYTEELRHVAPGLHLELRGPOL 379
QY 367 WARMGKCKYWEVGGPPGSGASSTPACLLPRNCPTPIFDRVFFQELVEFRARORRGP 426
DB 380 WARMGKCKYWEVGGPPGSGASSTPACLLPRNCPTPIFDRVFFQELVEFRARORRGP 439
QY 427 RYTIYLGFGQDLSAGRPKEKSLVLEPWLCRVHLEGTQREGVSSLDSSDLCLSSAN 486
DB 440 RYTIYLGFGQDLSAGRPKEKSLVLEPWLCRVHLEGTQREGVSSLDSSDLCLSSAN 499
QY 487 SLYDDIECFMLELEQPA 503
DB 500 SLYDDIECFMLELEQPA 516
RESULT 11
ADP24098
ID ADP24098 standard; protein; 516 AA.
XX
XX AC ADP24098;
XX
XX 18-NOV-2004 (first entry)
XX
XX PRO polypeptide SEQ ID NO:1276.
XX
XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
XX osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
XX antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX Unidentified.
XX
XX WO2004041170-A2.
XX
XX 21-MAY-2004.
XX
XX 30-OCT-2003; 2003WO-US034312.
XX
XX 01-NOV-2002; 2002US-0423394P.
XX (GETH) GENENTECH INC.
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WJ;
PI Wu TD;
XX WPI; 2004-419628/39.
DB N-PSDB; ADP24097.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 1276; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antitumoric, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
SQ Sequence 516 AA;

Query Match 98.2%; Score 2682; DB 8; Length 516;
Best Local Similarity 99.2%; Pred. No. 5.9e-193;
Matches 493; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RAAPRVLFGEWLLGESSGCGYEGQLWLDDEARTCFRVPWKHFARKDSEADARIFKAWAVA 66
DB 20 RAAPRVLFGEWLLGESSGCGYEGQLWLDDEARTCFRVPWKHFARKDSEADARIFKAWAVA 79
QY 67 RGRWPPSSRGSGGPPPEAETAEAGWKTNFRCALSTRFVFLMDNSGDDPADPHKVVALS 126
DB 80 RGRWPPSSRGSGGPPPEAETAEAGWKTNFRCALSTRFVFLMDNSGDDPADPHKVVALS 139
QY 127 ELCWREGPGTDTAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDKGDLLLQA 186
DB 140 ELCWREGPGTDTAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDEGDLLLQA 199
QY 187 VQOSCLADHLLTASWGADVPVTKAPGEGGGLPLTGACAGGPGLPAGELYGWAVETTPSP 246
DB 200 VQOSCLADHLLTASWGADVPVTKAPGEGGGLPLTGACAGGPGLPAGELYGWAVETTPSP 259
QY 247 GPQPAALTTEAAAPSPHQAEPYLSPSACTAVQSPSGALDVTIMYKGRVTLQKVVG 306
DB 260 GPQPAALTTEAAAPSPHQAEPYLSPSACTAVQSPSGALDVTIMYKGRVTLQKVVG 319
QY 307 HPSCTFLYGPDPVAVRATDPPQVAFPSPAELPDQKQLRYTEELRHVAPGLHLELRGPQL 366
DB 320 HPSCTFLYGPDPVAVRATDPPQVAFPSPAELPDQKQLRYTEELRHVAPGLHLELRGPQL 379
QY 367 WARMGCKVYWEVGGPGSASPTACLLPRNCDTPIFDRFVFFQELVEFRARQRGSP 426
DB 380 WARMGCKVYWEVGGPGSASPTACLLPRNCDTPIFDRFVFFQELVEFRARQRGSP 439
QY 427 RYTIYLGFGODLSAGRPKEKSLVVLKLEPWLCHVHLEGTQREGVSSLDSDLDLCLSSAN 486
DB 440 RYTIYLGFGODLSAGRPKEKSLVVLKLEPWLCHVHLEGTQREGVSSLDSDLSLCLSSAN 499

QY 487 SLYDDIECFLEMEQPA 503
DB 500 SLYDDIECFLEMEQPA 516

RESULT 12

ADJ75411
ID ADJ75411 standard; protein; 474 AA.

XX AC ADJ75411;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene related amino acid sequence SEQ ID NO:663.

XX KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.

XX OS Homo sapiens.

XX PN EP1394274-A2.

XX PD 03-MAR-2004.

XX PD 04-AUG-2003; 2003EP-00254857.

XX PF 06-AUG-2002; 2002JP-00229312.

XX PR 20-MAR-2003; 2003JP-00077212.

XX PA (GENO-) GENOX RES INC.

XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 663; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

QY 421 ORRGSPRYTILGFGODLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSDL 480
 Db 392 ORRGSPRYTILGFGODLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSSL 451
 QY 481 CLSSANSLYDDIECFLEMELEQPA 503
 Db 452 CLSSANSLYDDIECFLEMELEQPA 474

RESULT 14
 ABM83030
 ID ABM83030 standard; protein; 441 AA.
 AC ABM83030;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:3279.
 XX
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 XX
 OS Homo sapiens.
 XX
 PN WO2004023973-A2.
 XX
 PD 25-MAR-2004.
 XX
 PF 12-SEP-2003; 2003WO-US028227.
 XX
 PR 12-SEP-2002; 2002US-0410259P.
 PR 12-SEP-2002; 2002US-0410260P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patury S, Shi X, Suarez CJ;
 XX
 DR WPI; 2004-329368/30.
 DR N-PSDB; ACN41682.
 XX
 PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 PS Claim 27; Page; 190pp; English.
 XX
 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for sonatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIFO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 441 AA;

Query Match

81.5%; Score 2225.5; DB 8; Length 441;

Best Local Similarity 84.3%; Pred. No. 9.8e-159;
 Matches 419; Conservative 0; Mismatches 3; Indels 75; Gaps 1;
 QY 7 RAAPRVLFGEWLLGETSSCYEGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVA 66
 Db 20 RAAPRVLFGEWLLGETSSCYEGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVA 79
 QY 67 RGRWPPSSRGGGPPPEAETAERAGWKTNPRCALRSTRRRFVMLRDNDSGDPADPHKVYALSR 126
 Db 80 RGRWPPSSRGGGPPPEAETAERAGWKTNPRCALRSTRRRFVMLRDNDSGDPADPHKVYALSR 139
 QY 127 ELCWREGPGTQDTEAEAPAAVPPQGGPPGPFPLATHAGLQAPGPLPAGADKGDLLLOA 186
 Db 140 ELCWREGPGTQDTEAEAPAAVPPP-----163
 QY 187 VQSCCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 246
 Db 164 -----XGFCPLPAGELYGWAVETTPSP 184
 QY 247 GPQPAALTGEEAAAPESPHQAPSYLSPSACTAVQEPSFGALDVTIMYKGRTVLQKXVG 306
 Db 185 GPQPAALTGEEAAAPESPHQAPSYLSPSACTAVQEPSFGALDVTIMYKGRTVLQKXVG 244
 QY 307 HPSCTFLYGPDPVAVRATDPOQVAPSPAEPLDQKOLRYTELLRHVAPGLHLELGPOL 366
 Db 245 HPSCTFLYGPDPVAVRATDPOQVAPSPAEPLDQKOLRYTELLRHVAPGLHLELGPOL 304
 QY 367 WARRMGKCKVYWEVGGPPGSGASPTPACLLPNCOTPIEDFRVFFOELVEFRARQRGSP 426
 Db 305 WARRMGKCKVYWEVGGPPGSGASPTPACLLPNCOTPIEDFRVFFOELVEFRARQRGSP 364
 QY 427 RYTIYLGFGQDLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSDLCLSSAN 486
 Db 365 RYTIYLGFGQDLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSSLCLSSAN 424
 QY 487 SLYDDIECFLEMELEQPA 503
 Db 425 SLYDDIECFLEMELEQPA 441

RESULT 15
 ABM83031
 ID ABM83031 standard; protein; 440 AA.
 AC ABM83031;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:3280.
 XX
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 XX
 OS Homo sapiens.
 XX
 PN WO2004023973-A2.
 XX
 PD 25-MAR-2004.
 XX
 PF 12-SEP-2003; 2003WO-US028227.
 XX
 PR 12-SEP-2002; 2002US-0410259P.
 PR 12-SEP-2002; 2002US-0410260P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patury S, Shi X, Suarez CJ;

XX WPI: 2004-329368/30.
DR N-PSDB; ACN41683.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorders, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 440 AA;
Query Match 80.9%; Score 2210; DB 8; Length 440;
Best Local Similarity 84.1%; Pred. No. 1.4e-157;
Matches 418; Conservative 0; Mismatches 3; Indels 76; Gaps 2;
Qy 7 RAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIFKAWAVA 66
Db 20 RAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIFKAWAVA 79
Qy 67 RGRWPPSSRGPPPEAETRAERAGWKNFRCAKSTRFRVWLRDNSGDPADPHKVVALSR 126
Db 80 RGRWPPSSRGPPPEAETRAERAGWKNFRCAKSTRFRVWLRDNSGDPADPHKVVALSR 139
Qy 127 ELCWREGFGTDQTEAEAPAAVPPPGGPGFPLAHTAGLQAPGPLPAPAGDKGDLQLQA 186
Db 140 ELCWREGFGTDQTEAEAPAAVPP----- 163
Qy 187 VQOSCLADHLLTASWAGADPVTTKAPGEGQEGELPLTGACAGPGPLPAGELYGWAVETTPSP 246
Db 164 -----XGPGLPAGELYGWAVETTPSP 184
Qy 247 GPQPAALTGEEAAPESPQAEPLSPSACTAVQEPSGALDVTIMYKGRVTLOKVG 306
Db 185 GPQPAALTGEEAAPESPQAEPLSPSACTAVQEPSGALDVTIMYKGRVTLOKVG 244
Qy 307 HPSCTFLYGPDPVAVRATDQQVAFPPSPAEPLDQKQLRYTEELRHVAPGLHLELRGPQL 366
Db 245 HPSCTFLYGPDPVAVRATDQQVAFPPSPAEPLDQKQLRYTEELRHVAPGLHLELRGPQL 304
Qy 367 WARMGKCKVYWEVGGPPGASPTACLLPRNCDDTIFDPRVPFQELVEFRAPQRGSP 426
Db 305 WARMGKCKVYWEVGGPPGASPTACLLPRNCDDTIFDPRVPFQ-LVEFRAPQRGSP 363
Qy 427 RYTIYLGFGDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDLDLCLSSAN 486
Db 364 RYTIYLGFGDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDLDLCLSSAN 423
Qy 487 SLYDDIECFLEMELEQPA 503
Db 424 SLYDDIECFLEMELEQPA 440

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:40:43 ; Search time 44 Seconds
(without alignments)
853.375 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAAAPVLFGEWLLG.....SANSLYDDIECFMELSQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/ACTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2715	99.4	580	4	US-09-949-016-11716
2	2695	98.7	503	3	US-08-999-774A-10
3	2686	98.4	538	4	US-09-949-016-11715
4	2529.5	92.6	551	4	US-09-949-016-8970
5	1479	54.2	288	4	US-09-949-016-8969
6	458	16.8	427	4	US-09-949-016-6096
7	458	16.8	427	4	US-09-949-016-7374
8	454	16.6	427	3	US-08-705-771-16
9	454	16.6	427	4	US-09-417-540-16
10	376	13.8	450	2	US-08-611-280-2
11	376	13.8	450	3	US-09-195-940-2
12	376	13.8	450	3	US-09-562-466-2
13	363	13.3	486	4	US-09-949-016-7334
14	361.5	13.2	450	2	US-08-611-280-25
15	361.5	13.2	450	3	US-09-195-940-25
16	361.5	13.2	450	3	US-09-562-466-25
17	357	13.1	451	3	US-08-654-482-14
18	357	13.1	451	4	US-09-949-016-6229
19	328.5	12.0	439	3	US-09-230-371A-29
20	308.5	11.3	439	4	US-09-949-016-11483
21	306	11.2	425	3	US-09-230-371A-30
22	270.5	9.9	379	4	US-09-079-030-81
23	259.5	9.5	373	4	US-09-149-476-374
24	228	8.3	107	3	US-08-654-482-5
25	222.5	8.1	108	3	US-08-654-482-1
26	222.5	8.1	108	3	US-08-654-482-2
27	218	8.0	107	3	US-08-654-482-7

28	214	7.8	136	4	US-09-513-999C-5424	Sequence 5424, Ap
29	212	7.8	107	3	US-08-654-482-6	Sequence 6, Appli
30	198.5	7.3	102	4	US-09-513-999C-5034	Sequence 5034, Ap
31	176	6.4	325	4	US-09-949-016-6188	Sequence 6188, Ap
32	176	6.4	390	4	US-09-949-016-7930	Sequence 7930, Ap
33	169.5	6.2	802	4	US-09-252-991A-25050	Sequence 25050, A
34	167	6.1	352	4	US-09-949-016-11359	Sequence 11359, A
35	166	6.1	108	3	US-08-654-482-4	Sequence 4, Appli
36	161.5	5.9	108	3	US-08-654-482-3	Sequence 3, Appli
37	157.5	5.8	705	4	US-09-538-092-945	Sequence 945, App
38	156.5	5.7	420	4	US-09-902-540-13993	Sequence 13993, A
39	155.5	5.7	713	4	US-09-949-016-9700	Sequence 9700, Ap
40	155	5.7	561	1	US-08-642-255-52	Sequence 52, Appli
41	154.5	5.7	1064	1	US-08-642-255-62	Sequence 52, Appli
42	153.5	5.6	960	3	US-09-219-849-5	Sequence 5, Appli
43	153	5.6	349	4	US-09-919-497-76	Sequence 76, Appli
44	152	5.6	355	3	US-08-483-533-41	Sequence 41, Appli
45	152	5.6	355	3	US-09-283-471A-41	Sequence 41, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-11716
; Sequence 11716, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11716
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11716

Query Match	99.4%	Score	2715	DB	4	Length	580
Best Local Similarity	99.4%	Pred. No.	1.8e-224				
Matches	500	Conservative	1	Mismatches	2	Indels	0
Gaps	0						
QY	1	MALAPERAAAPVLFGEWLLG	ISSGCGEQLWLDARTCFRVPWKHFARKDLSEADARIF	60			
Db	78	MALAPERAAAPVLFGEWLLG	ISSGCGEQLWLDARTCFRVPWKHFARKDLSEADARIF	137			
QY	61	KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTFRFCALRSTRFVMLRDNSGDPADPHK	120				
Db	138	KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTFRFCALRSTRFVMLRDNSGDPADPHK	197				
QY	121	VVALSRELWCWRGPGTDTQTEAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDKG	180				
Db	198	VVALSRELWCWRGPGTDTQTEAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDKG	257				
QY	181	DLLLOAVQOCSLADHLLLTASWGCADVPVTKAPGEGGLPLTGCACAGGGLPAGELGWAV	240				
Db	258	DLLLOAVQOCSLADHLLLTASWGCADVPVTKAPGEGGLPLTGCACAGGGLPAGELGWAV	317				
QY	241	ETTPSGPGOPAAATTGEEAAAPSPHQAEPYLSFSPSACTAVQSPFGALDVTIMYKRTV	300				
Db	318	ETTPSGPGOPAAATTGEEAAAPSPHQAEPYLSFSPSACTAVQSPFGALDVTIMYKRTV	377				
QY	301	LOKVGVHPSCTFLYGPDPVAVRATDQQVAFSPSPAEPLDQKQLRYTEELLRHVAPGLHLE	360				

Db 378 LQKVGHPSTCTFLYGPDPFAVRATDPQVAFSPPAELPDQKQLRYTEELLRHVAPGLHLE 437
Qy 361 LRGPQLWARRMGKCKVYWEVGGPGSASPTTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
Db 438 LRGPQLWARRMGKCKVYWEVGGPGSASPTTACLLPRNCDTPIFDFRVFFQELVEFRAR 497
Qy 421 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDLDL 480
Db 498 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDLSL 557
Qy 481 CLSSANSLYDDIECFLEMELEQPA 503
Db 558 CLSSANSLYDDIECFLEMELEQPA 580
RESULT 2
US-08-999-774A-10
; Sequence 10, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,774A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,818
; FILING DATE: 11-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-999-774A-10
Query Match 98.7%; Score 2695; DB 3; Length 503;
Best Local Similarity 99.0%; Pred. No. 7.6e-223;
Matches 498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MALAPERAAARVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIF 60
Db 1 MALAPERAAARVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIF 60
Qy 61 KAWAVARGWPPSSRGSGPPPEAETAAERAGWKTNFRCALRSTRFRVMLRDNNSGDPADPHK 120

Db 61 KAWAVARGWPPSSRGSGPPPEAETAAERAGWKTNFRCALRSTRFRVMLRDNNSGDPADPHK 120
Qy 121 VVALSRELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
Db 121 VVALSRELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
Qy 181 DLLQVQVQSCVLADHLLTASWGADPVPTTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
Db 181 DLLQVQVQSCVLADHLLTASWGADPVPTTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
Qy 241 ETTSPSPGPPAALTTCGAAAPSPHQAEPYLSPPSACTAVQPPSPGALDVTIMYKRTV 300
Db 241 EKTSPSPGPPAALTTCGAAAPSPHQAEPYLSPPSACTAVQPPSPGALDVTIMYKRTV 300
Qy 301 LQKVVGHPSTCTFLYGPDPFAVRATDPQVAFSPPAELPDQKQLRYTEELLRHVAPGLHLE 360
Db 301 LQKVVGHPSTCTFLYGPDPFAVRATDPQVAFSPPAELPDQKQLRYTEELLRHVAPGLHLE 360
Qy 361 LRGPQLWARRMGKCKVYWEVGGPGSASPTTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
Db 361 LRGPQLWARRMGKCKVYWEVGGPGSASPTTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
Qy 421 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDLDL 480
Db 421 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDLSL 480
Qy 481 CLSSANSLYDDIECFLEMELEQPA 503
Db 481 CLSSANSLYDDIECFLEMELEQPA 503
RESULT 3
US-09-949-016-11715
; Sequence 11715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11715
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11715
Query Match 98.4%; Score 2686; DB 4; Length 538;
Best Local Similarity 99.4%; Pred. No. 4.9e-222;
Matches 494; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 7 RAAPRVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIFKAWAVA 66
Db 42 RAAPRVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIFKAWAVA 101
Qy 67 RGRWPPSSRGSGPPPEAETAAERAGWKTNFRCALRSTRFRVMLRDNNSGDPADPHKVVALSR 126
Db 102 RGRWPPSSRGSGPPPEAETAAERAGWKTNFRCALRSTRFRVMLRDNNSGDPADPHKVVALSR 161
Qy 127 ELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKGDLQLQA 186
Db 162 ELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKGDLQLQA 221
Qy 187 VQOSCLADHLLTASWGADPVPTTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 246

Db 222 VQSCADLLTASMGADVPVTKAPGEGGGLPLTGACAGGGLPAGELYGNVAVETTPSP 281
QY 247 GQPAALTTGEEAAAPESPHQAEPIYSPSPSACTAVQEPSFGALDVTIMYKGRITVLQKVVG 306
Db 282 GQPAALTTGEEAAAPESPHQAEPIYSPSACTAVQEPSFGALDVTIMYKGRITVLQKVVG 341
QY 307 HPSCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKOLRYTELLRHVAPGLHLELRGPOL 366
Db 342 HPSCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKOLRYTELLRHVAPGLHLELRGPOL 401
QY 367 WARRMGCKVYVEVGGPGSASPSTACLLPRNCDTPIFDFRVFRELVEFRARQRGSP 426
Db 402 WARRMGCKVYVEVGGPGSASPSTACLLPRNCDTPIFDFRVFRELVEFRARQRGSP 461
QY 427 RYTIYLGFGQDLSAGRPKESKLVLVKLEPWLCRVHLEGTQREGVSSLDSSLSLCLSSAN 486
Db 462 RYTIYLGFGQDLSAGRPKESKLVLVKLEPWLCRVHLEGTQREGVSSLDSSLSLCLSSAN 521
QY 487 SLYDDIECFMELEQPA 503
Db 522 SLYDDIECFMELEQPA 538

RESULT 4

US-09-949-016-8970
; Sequence 8970, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8970
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8970

Query Match 92.6%; Score 2529.5; DB 4; Length 551;
Best Local Similarity 93.6%; Pred. No. 1.3e-208;
Matches 471; Conservative 1; Mismatches 2; Indels 29; Gaps 1;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCVGLOWLDEARTCFRPWKHPARKDLSEADARIF 60
Db 78 MALAPERAAPRVLFGEWLLGEISSGCVGLOWLDEARTCFRPWKHPARKDLSEADARIF 137
QY 61 KAWAVARGWPPSSRGSGPPPEAETARAGKTNFRCAIRSTRFRVWLNRDNGDPADPHK 120
Db 138 KAWAVARGWPPSSRGSGPPPEAETARAGKTNFRCAIRSTRFRVWLNRDNGDPADPHK 197
QY 121 VVALSRELWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGFLPAPAGDXG 180
Db 198 VVALSRELWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGFLPAPAGDXG 257
QY 181 DLLLQVQSCADHLLTASWGADPVTAKAPGQSGGLPLTGACAGGGLPAGELYGNVAV 240
Db 258 DLLLQVQSCADHLLTASWGADPVTAKAPGQSGGLPLTGACAGGGLPAGELYGNVAV 303
QY 241 ETTTPSPQPAALTTGEEAAAPESPHQAEPIYSPSPSACTAVQEPSFGALDVTIMYKGRV 300
Db 304 -----GEEAAAPESPHQAEPIYSPSPSACTAVQEPSFGALDVTIMYKGRV 348

QY 301 LOKVVGHPSCCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKOLRYTELLRHVAPGLHLE 360
Db 349 LOKVVGHPSCCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKOLRYTELLRHVAPGLHLE 408
QY 361 LRGPOLWARRMGCKVYVEVGGPGSASPSTACLLPRNCDTPIFDFRVFRELVEFRAR 420
Db 409 LRGPOLWARRMGCKVYVEVGGPGSASPSTACLLPRNCDTPIFDFRVFRELVEFRAR 468
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKESKLVLVKLEPWLCRVHLEGTQREGVSSLDSSLSL 480
Db 469 QRRGSPRYTIYLGFGQDLSAGRPKESKLVLVKLEPWLCRVHLEGTQREGVSSLDSSLSL 528
QY 481 CLSSANSLYDDIECFMELEQPA 503
Db 529 CLSSANSLYDDIECFMELEQPA 551

RESULT 5

US-09-949-016-8969
; Sequence 8969, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8969
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8969

Query Match 54.2%; Score 1479; DB 4; Length 288;
Best Local Similarity 95.2%; Pred. No. 7.3e-119;
Matches 278; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 212 GREGQGLPLTGACAGGGLPAGELYGNVAVETTPSPGQPAAALTTGEEAAAPESPHQAEPIYL 271
Db 1 GRGPRSCPTT-----TGGLPAGELYGNVAVETTPSPGQPAAALTTGEEAAAPESPHQAEPIYL 56
QY 272 SPSPSACTAVQEPSFGALDVTIMYKGRITVLQKVGHPSCTFLYGPDPVAVRATDPOQVAF 331
Db 57 SPSPSACTAVQEPSFGALDVTIMYKGRITVLQKVGHPSCTFLYGPDPVAVRATDPOQVAF 116
QY 332 PSFAELPDQKOLRYTELLRHVAPGLHLELRGPOLWARRMGCKVYVEVGGPGSASPST 391
Db 117 PSFAELPDQKOLRYTELLRHVAPGLHLELRGPOLWARRMGCKVYVEVGGPGSASPST 176
QY 392 PACLLPRNCDTPIFDFRVFRELVEFRARQRGSPRYTIYLGFGQDLSAGRPKESKLVLV 451
Db 177 PACLLPRNCDTPIFDFRVFRELVEFRARQRGSPRYTIYLGFGQDLSAGRPKESKLVLV 236
QY 452 KLEPWLCRVHLEGTQREGVSSLDSSLSLCLSSANSYDDIECFMELEQPA 503
Db 237 KLEPWLCRVHLEGTQREGVSSLDSSLSLCLSSANSYDDIECFMELEQPA 288

RESULT 6

US-09-949-016-6096
; Sequence 6096, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

QY 301 LQKVGH-P-SCFTLYGPPDPVAVRATDPOQVAPSPAEPLDQKOLRYTELLRHVAGLHL 359
Db 264 KELLTSSPEGCRISHG---HTYDVSNLDQVLFYP---DNGQRKNIEKLLSHLERGLVL 317
QY 360 ELRGPOLWARRMGCKVYWEVGGPPGAS--PSTPACLLPRNCDTPIFDFRVFFOELVEF 417
Db 318 WMAPDGLYAKLCOSRIYD--GPLALCSDRNK-----LERDOTCKLFTQQLSELQVF 371
QY 418 RARQRGSPRYTYILGFGQDLSAGRPKEKSLVLVLEPWLCL 459
Db 372 -AAHGRPAFRFQVTLFCGEEFPDPQ-RQRKLTAHVPELLAR 411

RESULT 12
US-09-562-466-2
; Sequence 2, Application US/09562466
; Patent No. 6369202
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; Grossman, Alex
; Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/562,466
; FILING DATE: 01-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/195,940
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-562-466-2

Query Match 13.8%; Score 376; DB 3; Length 450;
Best Local Similarity 28.1%; Pred. No. 5.4e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIPKAVAVARGRPPSS 74
Db 26 QWLIDQIDSGKYPGLVWENEKSVFRIPWKHAGKQDYNEEDAALFKAWALFKGKF--- 81
QY 75 RGGGPPPEAETAERAGKTNFCALRSTRFVMLRDSG-DPADPHKVVALSRELWCWREG 133
Db 82 REGIDKPDPT-----WKTLRCALNKNSDFEELVERSQLDISDPYKVVRI----- 127
QY 134 PGTDQTEAEAPAAVPPPGGPPGPPFLAHTHAGLQAPGP-LPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKGAKQLTLDQTQWANGHPYPTAP---YGSLEPAQVHNY 169
QY 191 CLADHLLTASWGADPVPTKAPGEGQLPLTGACG-----GPGLP-----AGELYGWAVE 241

RESULT 13

US-09-949-016-7334
; Sequence 7334, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7334
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7334

Query Match 13.3%; Score 363; DB 4; Length 486;

Best Local Similarity 27.5%; Pred. No. 7.9e-23;

Matches 136; Conservative 73; Mismatches 193; Indels 92; Gaps 24;

QY 16 EWLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIPKAVAVARGRPPSS 74

Db 61 QWLIDQIDSGKYPGLVWENEKSVFRIPWKHAGKQDYNEEDAALFKAWALFKGKF--- 116

QY 75 RGGGPPPEAETAERAGKTNFCALRSTRFVMLRDSG-DPADPHKVVALSRELWCWREG 133

Db 117 REGIDKPDPT-----WKTLRCALNKNSDFEELVERSQLDISDPYKVVRIPEGAK-- 169

QY 134 PGTDQTEAEAPAAVPPPGGPPGPPFLAHTHAGLQAPGP-LPAPAGDKGDLILLQAVQOSCL 192

Db 170 -GAKQLTLEDPOW-----SMSEHYT-WTTPYPSLPA-----QOVHNYM 206

QY 193 ADHLITASWGADPVPTKAPGEGQLPLT-----GACAGGGLPAGELYGWAVET 242

Db 207 PE--LDRSW-RDYVPDQHPPEIPYQCPMTFGPRGHHWQGPACENGQV-TGTFYACAPPE 262

QY 243 TSPG-PQPAALTTEAAAPESPHQAEPLSPSACTAVQSPGALDVTIMYKGRVL 301

Db 263 SQAGVPTSPRSAEALA-----FSDC-----RLHICLYRIILVK 299

QY 302 QKVVGHP-SCFTLYGPPDPVAVRATDPOQVAPSPAEPLDQKOLRYTELLRHVAPGLHLE 360

Db 300 ELTSSPEGCRISHG---HTYDVSNLDQVLFYP---EDNGQRKNIEKLLSHLERGVWLW 353

Search completed: May 6, 2005, 16:49:19
Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:44:59 ; Search time 58 Seconds
(without alignments)
2893.107 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSYDDIECFMLELEQPA 503

Scoring table: BLASTSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
- 12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
- 17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832	30.5	164	16	US-10-755-889-356
2	458	16.8	427	17	US-10-741-600-1291
3	458	16.8	427	17	US-10-741-600-1293
4	454	16.6	427	9	US-09-975-253-2
5	454	16.6	427	12	US-09-975-253-2
6	454	16.6	427	15	US-10-453-478-16
7	446	16.3	175	15	US-10-264-049-4045
8	431	15.8	467	17	US-10-840-1398-2
9	431	15.8	467	17	US-10-840-1398-3
10	364.5	13.3	452	17	US-10-741-600-1292
11	337.5	12.4	497	17	US-10-492-043-2
12	335.5	12.3	72	15	US-10-276-774-1410
13	332.5	12.2	473	17	US-10-492-043-30

14	328.5	12.0	393	17	US-10-492-043-5	Sequence 5, Appli
15	320	11.7	392	15	US-10-276-774-2333	Sequence 2333, Ap
16	259.5	9.5	373	10	US-09-809-391-374	Sequence 373, App
17	259.5	9.5	373	10	US-09-882-171-374	Sequence 374, App
18	259.5	9.5	373	15	US-10-164-861-374	Sequence 374, App
19	188.5	6.9	625	16	US-10-437-963-165015	Sequence 165015,
20	176	6.4	292	9	US-09-805-020-46	Sequence 46, Appl
21	176	6.4	395	15	US-10-284-049-2929	Sequence 2929, Ap
22	171	6.3	411	15	US-10-108-260A-4496	Sequence 4496, Ap
23	167	6.1	1023	10	US-09-893-519A-14	Sequence 14, Appl
24	166.5	6.1	1014	17	US-10-901-816A-12	Sequence 12, Appl
25	162	5.9	577	15	US-10-055-569A-4	Sequence 4, Appli
26	161.5	5.9	551	15	US-10-055-569A-42	Sequence 42, Appl
27	157	5.7	1413	15	US-10-288-798-24	Sequence 24, Appl
28	157	5.7	1413	15	US-10-362-892-34	Sequence 24, Appl
29	156	5.7	696	15	US-10-092-900A-272	Sequence 272, App
30	155.5	5.7	668	14	US-10-240-154-20	Sequence 20, Appl
31	155	5.7	576	15	US-10-055-569A-40	Sequence 40, Appl
32	155	5.7	576	17	US-10-898-142-13	Sequence 13, Appl
33	154.5	5.7	409	16	US-10-437-963-140753	Sequence 140753,
34	153.5	5.6	960	15	US-10-342-331-5	Sequence 5, Appl
35	153	5.6	349	9	US-09-919-497-76	Sequence 76, Appl
36	153	5.6	542	15	US-10-276-774-1386	Sequence 1386, Ap
37	152	5.6	1466	15	US-10-402-089-4	Sequence 4, Appli
38	152	5.6	1466	15	US-10-402-089-6	Sequence 6, Appli
39	152	5.6	1466	15	US-10-402-072A-4	Sequence 4, Appli
40	152	5.6	1466	15	US-10-402-072A-6	Sequence 6, Appli
41	151	5.5	369	16	US-10-437-963-162013	Sequence 162013,
42	151	5.5	507	16	US-10-437-963-187017	Sequence 187017,
43	150.5	5.5	550	9	US-09-976-740-47	Sequence 47, Appl
44	150.5	5.5	550	13	US-10-023-529-47	Sequence 47, Appl
45	150.5	5.5	550	13	US-10-023-523-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-10-755-889-356
; Sequence 356, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 356
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-356

Query Match	30.5%	Score	832	DB	16	Length	164
Best Local Similarity	100.0%	Pred. No.	1e-50				
Matches	151	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MALAPERAPRVLFGEWLLGESSCYEGLWLDARTCFRVPWKHFARKDSEADARIF	60				
Db	1	MALAPERAPRVLFGEWLLGESSCYEGLWLDARTCFRVPWKHFARKDSEADARIF	60				
Qy	61	KAWAVARGWPPSSRGCGPPPEAETAEARAGWKTNFRFCALRSTRFRFVLMRDNSGDPADPHK	120				
Db	61	KAWAVARGWPPSSRGCGPPPEAETAEARAGWKTNFRFCALRSTRFRFVLMRDNSGDPADPHK	120				
Qy	121	VVALSRELCWREGPGTDTQTEAEAPAAVPPQ	151				

OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val
US-09-975-253-2

Query Match 16.6%; Score 454; DB 9; Length 427;
Best Local Similarity 28.0%; Pred. No. 1.1e-23;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;
10 PRVLFGEWLLGEISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARGR 69
6 PRXL--PMLVSQLDLQLEGVAVWVKSRTRFRIPWKHGLRQDAQOEDFGIFQAWAEATGA 63
70 WPPSSRGSGPPPEAEATERRAGWKTFRFCALRSTRFRVMLRDNSSGDPADPHKVYALSREL 129
64 YVP---GRDKDLP-----WKFRSALNRKEGLRLAEDRSKDPHDPHKIYEFV--- 110
130 WREGPTDQTEABEAPAAVPPPGGPPGPFPLAHTAGLQAPGFLPAPAGDKGDLQLLQAVQ 189
111 -NSGVG---DFSQDPTSDTNGG-----GSTSD-----TQ 136
190 SCLADHLLTASGADVPVTKAPGEGQEGGLPLTGACAGGGLPAGELYGWAVETTPSPGPQ 249
137 EDILDELL--GNMVLAPLP-----DFGPP-----SLAVAPEPCPQ 169
250 PAALTGEEAAAPESPHQAEPLYSPSPSACTAVOEP--SPGALDVTIMYKGRTVLQKVVGHP 308
170 FLRSPSLNPTP-----FPNLGSENPLKRLVPGEEWEFEVTAFYRGQVFOQTISCP 223
309 SCTFLYGGPPDPAVRATDPOQVAFPSA--ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
224 EGLRLVG--SEVGDRTLPGWPVTLPGMSLTDGVMYSYVRHVLSCLGGGLALWRAGQLW 282
368 ARMGCKKYWEVG-----GPPGSASPTPACLLPRNCPTPIFDFRVFQELVEPR 418
283 AORLGHCITYWAVSEELLPSNGHGPGE-----VPKXEGGVDFLGPFIVDLITFT 333
419 ARQRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 478
334 EGSGR--SPRYALWFCVGSWPQDQWTKRLVMKVVTCLRALVEMARVGGASSLENT-V 391
479 DLCLSSANSL---YDDIECFML 499
392 DLHISNSHPLSLTSDQYKAYLQDL 415

RESULT 5
US-09-975-253-2
Sequence 2, Application US/0975253
Publication No. US20050054033A9
GENERAL INFORMATION:
APPLICANT: Paul Moore et al.
TITLE OF INVENTION: Methods and Compositions for Treating and Preventing Infection
TITLE OF INVENTION: Using Human Interferon Regulatory Factor 3
FILE REFERENCE: PF196P1
CURRENT APPLICATION NUMBER: US/09/975,253
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/239,963
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 427
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8)
OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val
US-09-975-253-2

Query Match 16.6%; Score 454; DB 12; Length 427;
Best Local Similarity 28.0%; Pred. No. 1.1e-23;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;

10 PRVLFGEWLLGEISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARGR 69
6 PRXL--PMLVSQLDLQLEGVAVWVKSRTRFRIPWKHGLRQDAQOEDFGIFQAWAEATGA 63
70 WPPSSRGSGPPPEAEATERRAGWKTFRFCALRSTRFRVMLRDNSSGDPADPHKVYALSREL 129
64 YVP---GRDKDLP-----WKFRSALNRKEGLRLAEDRSKDPHDPHKIYEFV--- 110
130 WREGPTDQTEABEAPAAVPPPGGPPGPFPLAHTAGLQAPGFLPAPAGDKGDLQLLQAVQ 189
111 -NSGVG---DFSQDPTSDTNGG-----GSTSD-----TQ 136
190 SCLADHLLTASGADVPVTKAPGEGQEGGLPLTGACAGGGLPAGELYGWAVETTPSPGPQ 249
137 EDILDELL--GNMVLAPLP-----DFGPP-----SLAVAPEPCPQ 169
250 PAALTGEEAAAPESPHQAEPLYSPSPSACTAVOEP--SPGALDVTIMYKGRTVLQKVVGHP 308
170 FLRSPSLNPTP-----FPNLGSENPLKRLVPGEEWEFEVTAFYRGQVFOQTISCP 223
309 SCTFLYGGPPDPAVRATDPOQVAFPSA--ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
224 EGLRLVG--SEVGDRTLPGWPVTLPGMSLTDGVMYSYVRHVLSCLGGGLALWRAGQLW 282
368 ARMGCKKYWEVG-----GPPGSASPTPACLLPRNCPTPIFDFRVFQELVEPR 418
283 AORLGHCITYWAVSEELLPSNGHGPGE-----VPKXEGGVDFLGPFIVDLITFT 333
419 ARQRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 478
334 EGSGR--SPRYALWFCVGSWPQDQWTKRLVMKVVTCLRALVEMARVGGASSLENT-V 391
479 DLCLSSANSL---YDDIECFML 499
392 DLHISNSHPLSLTSDQYKAYLQDL 415
RESULT 6
US-10-453-478-16
Sequence 16, Application US/10453478
Publication No. US20030208043A1
GENERAL INFORMATION:
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
Jian Ni and Jing-Shan Hu
TITLE OF INVENTION: Human Genes, Sequences and
Expression Products
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/453,478
FILING DATE: 04-Jun-2003
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/705,771
FILING DATE: August 30, 1996
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-346 (PF196)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 973-994-1700

TELEFAX: 973-994-1744
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 427 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-453-478-16

Query Match 16.6%; Score 454; DB 15; Length 427;
Best Local Similarity 28.0%; Pred. No. 1.1e-23;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;
QY 10 PVLFGWLGELSSGCVGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARG 69
DB 6 PRL--PWLVSQDLQGLEGVAMVNSKRTFRIPWKHGLRQDAQQDFGIFQWAEATGA 63
QY 70 WPPSSRGGPPPEAETAAERAGWKTNFRCALRSTRFVMLRDNSTGDPADPHKVALSREL 129
DB 64 YVP---GRDKPDLPT-----WKRFPSALNRKGLRLAEDRSKDPHPKIVEFV--- 110
QY 130 WRGPGTDTQTEARAPAAVPPQGGPGPFPLATHAGLQAPGPLPAPAGDKGDLQLQAVQ 189
DB 111 -NSGVG---DFSQPDTSPDTNGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASWGADVPVTRKAPGEGQGLPLTGACAGGPGLPAGELYGMAVEITPSPGQ 249
DB 137 EDLDEL--GNVLAPLP-----DPGPP-----SLVAPEPCPQ 169
QY 250 PAALTGEEAAPSPHOAEPLSPSPSACTAVOEP-SPGALDVTIMYKGRVTLQKVVGHP 308
DB 170 PLRSPSLDNTP-----FNLGSENPPLKRLVPGGEWEFEVTA FYRGVQVQQTISCP 223
QY 309 SCTFLYPPDPVAVRTPDQVAPSPA-ELPDQKQLRYTEELRHVAPGLHLELRGPQLW 367
DB 224 EGLRLVG-SEVGDRTLPGWVPTLPDQMSLTDRGVMSYVRHLVSLCGGLALWRAGOWLW 282
QY 368 ARWGKCKVYWEVG-----GPGSASPSPTACLLPNCDDTPIEDRFVFEQELVEER 418
DB 283 AQLRGHCHTVAWSEELPNSHGPDGE-----VPNDKEGVDFDLGPFVVDLITFT 333
QY 419 ARQRGSPRTYIYLGFGQDLSAGRPKEKSLVLEPWLRCRVHLEGTQREGVSSLSDDL 478
DB 334 EGSR-SPRYALMFCVGSWPQDQWTKRLVMVKVPTCLRALVEMARVGGASSENT-V 391
QY 479 DLCLSSANSL---YDDIECFIMEL 499
DB 392 DLHISNHPLSLTSDQYKAYLQDL 415

RESULT 7
US-10-264-049-4045
Sequence 4045, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133PI
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: Patent In Ver. 3.1
SEQ ID NO 4045
LENGTH: 175
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE

LOCATION: (4)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (34)_FEATURE
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (90)_FEATURE
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (106)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (166)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4045

Query Match 16.3%; Score 446; DB 15; Length 175;
Best Local Similarity 55.1%; Pred. No. 1.4e-23;
Matches 102; Conservative 9; Mismatches 16; Indels 58; Gaps 8;
QY 334 PAELPDQKQLRYTEELRHVAPGLHLELRGPQLWARRMGKCKVYWEVGGPPGASSTPA 393
DB 2 PAXLPDQKQLRYTEELRHVAPGLHLELRGPQXWARRMGKCKVYWEVGGPPGASSTPA 61
QY 394 CLLPNCDDTPIEDRFVFFQ-----ELVEPRARQ 421
DB 62 CLLPNCDDTPIEDRFVFFQVQVRPLAWGKPGSLGKHLPCQVLXKXSPSELVEF--- 117
QY 422 RGSPRY-TIY-----LGFGQDLSAGRPKEK-----SLVLVLEPWLRCRVHLEGTQREG 469
DB 118 RHGAPWLPDLYHLPLGRAG--PVSLLGGPRRRPGCEAGTLVW-----PMHLEGTXXK- 167
QY 470 VSSLD 474
DB 168 VGFLD 172

RESULT 8
US-10-840-139B-2
Sequence 2, Application US/10840139B
Publication No. US2005008985A1
GENERAL INFORMATION:
APPLICANT: Schutte, Brian C.
APPLICANT: Murray, Jeffrey C.
APPLICANT: Kondo, Shinji
APPLICANT: Dixon, Michael J.
TITLE OF INVENTION: IRF6 POLYMORPHISMS ASSOCIATED WITH CLEFT LIP AND/OR PALATE
FILE REFERENCE: P06215US01
CURRENT APPLICATION NUMBER: US/10/840,139B
CURRENT FILING DATE: 2004-05-06
PRIOR APPLICATION NUMBER: US 60/468,191
PRIOR FILING DATE: 2003-05-06
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent in version 3.3
SEQ ID NO 2
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
US-10-840-139B-2

Query Match 15.8%; Score 431; DB 17; Length 467;
Best Local Similarity 27.7%; Pred. No. 4.9e-22;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;
QY 1 MALAPERAPRVLFGWLLGEISSGCVGLQWLDEARTCFRVPWKHFAR-KDLSEADARI 59
DB 1 MALHPR-----RVRLKPLVLAQVDSGLYPGLIWLHRSKRFOIPWKATRHSPQEEENTI 56
QY 60 FKAWAVARGRWPSPSRGGPPPEAETAAERAGWKTNFRCALRSTRFVMLRDNSTGDPADP 118

```
Db 57 FKAWAVETGKY---QEGVDDPPAK-----WKAQLRCALNKSRFENLMDYGTKEVPMNP 107
Qy 119 HKVYALSRCLWREGPGTQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGD 178
Db 108 VKIY---QVC-----DIPQGG-----SIINPGSTGSAPWD 135
Qy 179 KGDLLQAVQOCLADHLHTAGWADPVETKAPGEQGLPL-----TGACAGGPGGLPAGE 234
Db 136 EXD---NDVDEEDEDE-LDQSHHVPIQDTFFFLNINGSMPAPASVGNCSVGNCSPEAV 191
Qy 235 LYCWAVETTPSPQCPQPAALTTGEAAPESPHQAEVLSPPSACTAVQEPSPGALDVTIM 294
Db 192 ---W-----PKTEPLEMEVPO--APIQFFYSSPELWISSLPMT-----DLDIKFO 231
Qy 295 YKGRTVLQKV-VGHP--SCTFLYGPDP-----AVRATDPQVAPFPAELPDQQLRYT 346
Db 232 YRGKEVGQMTVSNPGOCRLFGDGLGMPDQBELFGPVSLQGVKPGPHITNEKQLFT 291
Qy 347 BELLRHVAPGLHLELRGPOLWARRMKCKVYWEVGGPGSPSTPACILLPRNCOTPIPD 406
Db 292 SKLLDVMDRGLILEVSHAIYAIRLCQCKVYWSGPCAPSLVAPN---LIERQKKVKLPC 347
Qy 407 FRVFFOELVEFRARQRGRSPRTIYLGFGODLSAGRPKESLVLYKLEPWLGRVHLEGTQ 466
Db 348 LETFLSDLIAHQKGQIEKQPFPEIYLCFGEWPDGKPLERKILVQVIVPVARMIYEMPS 407
Qy 467 REGVSLDSDLDLCLSSANSLYDDIECFMEL 499
Db 408 GDFTRSFDSGVRQLQISTPD-TKDNIVAQLKQL 439

RESULT 9
US-10-840-139B-3
; Sequence 3, Application US/10840139B
; Publication No. US20050089885A1
; GENERAL INFORMATION:
; APPLICANT: Schutte, Brian C.
; APPLICANT: Murray, Jeffrey C.
; APPLICANT: Kondo, Shinji
; APPLICANT: Dixon, Michael J.
; TITLE OF INVENTION: IRF6 POLYMORPHISMS ASSOCIATED WITH CLEFT LIP AND/OR PALATE
; FILE REFERENCE: P06215US01
; CURRENT APPLICATION NUMBER: US/10/840,139B
; CURRENT FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: US 60/468,191
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: 5' UTR to Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(1)
; OTHER INFORMATION: Metille
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (2)..(2)
; OTHER INFORMATION: Ala2Val
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (6)..(6)
; OTHER INFORMATION: Arg6fs
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (17)..(17)
; OTHER INFORMATION: Gln17fs
; FEATURE:
```

```
; NAME/KEY: VARIANT
; LOCATION: (18)..(18)
; OTHER INFORMATION: Val18Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (18)..(18)
; OTHER INFORMATION: Val18Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (39)..(39)
; OTHER INFORMATION: Pro39Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (60)..(60)
; OTHER INFORMATION: Trp60Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (61)..(61)
; OTHER INFORMATION: Ala61Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (66)..(66)
; OTHER INFORMATION: Lys66Thr
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (68)..(68)
; OTHER INFORMATION: Gln68X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (70)..(70)
; OTHER INFORMATION: Gly70Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (76)..(76)
; OTHER INFORMATION: Pro76Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (82)..(82)
; OTHER INFORMATION: Gln82Lys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (84)..(84)
; OTHER INFORMATION: Arg84Cys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (84)..(84)
; OTHER INFORMATION: Arg84His
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (89)..(89)
; OTHER INFORMATION: Lys89Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (90)..(90)
; OTHER INFORMATION: Ser90Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (92)..(92)
; OTHER INFORMATION: Glu92X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (98)..(98)
; OTHER INFORMATION: Asp98His
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (118)..(118)
; OTHER INFORMATION: Gln118X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (156)..(156)
; OTHER INFORMATION: His156fs
; FEATURE:
; NAME/KEY: VARIANT
```

```
; LOCATION: (186)..(186)
; OTHER INFORMATION: Cys186X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (192)..(192)
; OTHER INFORMATION: Trp192X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (212)..(212)
; OTHER INFORMATION: Ser212fs
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (219)..(219)
; OTHER INFORMATION: Ser219fs
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (248)..(248)
; OTHER INFORMATION: Gly248fs
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (250)..(250)
; OTHER INFORMATION: Arg250Gln
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (253)..(253)
; OTHER INFORMATION: Tyr253X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (265)..(265)
; OTHER INFORMATION: Leu265fs
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (273)..(273)
; OTHER INFORMATION: Gln273Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (274)..(274)
; OTHER INFORMATION: Val274Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (294)..(294)
; OTHER INFORMATION: Leu294Pro
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (297)..(297)
; OTHER INFORMATION: Val297Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (320)..(320)
; OTHER INFORMATION: Lys320Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (321)..(321)
; OTHER INFORMATION: Val321Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (325)..(325)
; OTHER INFORMATION: Gly325Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (345)..(345)
; OTHER INFORMATION: Leu345Pro
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (347)..(347)
; OTHER INFORMATION: Cys347Phe
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (369)..(369)
; OTHER INFORMATION: Phe369Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (374)..(374)
```

```
; OTHER INFORMATION: Cys374Trp
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (388)..(388)
; OTHER INFORMATION: Lys388Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (393)..(393)
; OTHER INFORMATION: Gln393X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (412)..(412)
; OTHER INFORMATION: Arg412X
; FEATURE:
; NAME/KEY: VARIANT
```

Query Match 15.8%; Score 431; DB 17; Length 467;
Best Local Similarity 27.7%; Pred. No. 4.9e-22;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;

```
Qy 1 MALAPERAAPRVLFGEWLLGEISSGCGEGLWLDDEARTCFRVPWKHFAR-KDLSEADARI 59
Db 1 MALHPR----RVRLKPWLVAQVDSGLYPGLIWLHRDSKRFQIPWKHATRHSPOESENTI 56
Qy 60 FKAWAVARGRWPPSSRGGGPPPEAETAERAGWKTNFRCALRSTRFRVMLRDNISGD-PADP 118
Db 57 FKAWAVETGKY---QEGVDDDPDAK-----WKAQLRCALNKSRFENLMDYDGTKEVPMNP 107
Qy 119 HKVYALSRELWCWREGPGTQTEAEAAVPPPGGPPGLAHTHAGLOAPGLPAPAGD 178
Db 108 VKIY---QVC-----DIPQPG-----SIINPGSTGSAPWD 135
Qy 179 KGDLLLAQVQOSCLADHLLTASMGADPVPTKAPGEGOGLPL----TGACAGGPGLPAGE 234
Db 136 EKD---NDVDEEDEDE-LDQSHHVPIDQTFEFLNINSGMAPASVGNCSVCNCSPEAV 191
Qy 235 LYGMAVETTPSPGPQPAALTTGEAAAPESPHQAEPLYSLSPSPSACTAVQEPSGALDVIM 294
Db 192 ---W-----PKTEPLEMEVPO--APIQPFYSSPELWISLPMT-----DLDIKFQ 231
Qy 295 YKQRTVLQKV-VGHP-SCTFLVGPDP-----AVRATDQOVAFPSPAELPDKQLRYT 346
Db 232 YRGKEYGQTMVTSNPQGCRLFYGDGLGMPDQBELFGVSLQVFKPGPEHITNEKQKLT 291
Qy 347 EELLRHVAPGLHLELGPOLWARRMGKCKVYVGVGGPPGSGASPSPTACLLPRNCDTPIFD 406
Db 292 SKLLDVMDRGLILEVSGHAIYAIRLCQCKVYNSGPCAPSLVAPN----LIERQKVKLFC 347
Qy 407 FRVFFOELVEFRARQRRGSPRYTYIGFGODLSAGRPKEKSLVLVLEPMLCRVHLEGTQ 466
Db 348 LETFLSDLIAHQKGQIEKQPPFEIYLCFGEWPDGKPLERKLILVQIVPVVARMYEMFS 407
Qy 467 REGVSSILDSSDLLCLSSANSLYDDIECFMEL 499
Db 408 GDFTRSFDSGVSRLQISTPD-IKDNIVAQLKQL 439
```

```
RESULT 10
US-10-741-600-1292
; Sequence 1292, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1292
; LENGTH: 452
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
US-10-741-600-1292

Query Match      13.3%; Score 364.5; DB 17; Length 452;
Best Local Similarity 27.4%; Pred. No. 2.2e-17;
Matches 113; Conservative 43; Mismatches 153; Indels 103; Gaps 16;

QY 10 PRVLFGWLLGEISSCYEGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVARGR 69
DB 6 PRIL--PWLVSQDLGLQLEGVAVWVKSTRFRIPWKHGLRQDAQQBDFGIFQAWAEATGA 63
QY 70 WPPSSRGSGPPPEAEATERRAGWKTFRFCALRSTRFRVMLRDNSSGDPADPHKVVALSRELC 129
DB 64 YP-----GRDKPDLPT-----WKNRPSALNKKEGLRLAEDRSKDPDHPKHYEFV----- 110
QY 130 WREGPDTQTEAEAPAAVPPPGGPPGFLAHTAGLQAPGFLPAGDGKGLLLQAVQ 189
DB 111 -NSGVG---DFSQDPTSDTNGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASWGADVPVTKAPGEGGGLPLTGACAGGPGGLPAGELYGWAVETTPSPGPQ 249
DB 137 EDILDELL-GNMVLAPLP-----DPGPP-----SLAVAPPCPQ 169
QY 250 PAALTGGERAAEPSPHQAEPYLSPPSACTAVQEP--SPGALDVTIMYKGRTVLQKVVGHP 308
DB 170 PLRSPSLDNPT-----FNLGSENPLKRLLVGGEWEFEVTAFYRGQVFOQTISCP 223
QY 309 SCTFLYGGPDPAVRATDPOQVAPPSPA--ELPDQKQLRYTEELLRHVAPGLHLELRGPOLW 367
DB 224 EGLRVG-SEVGDRTPLGWPVTLDPGMSITDRGVMSYVRHVLSCLGGGLALWRAGQWLW 282
QY 368 ARMGCKVYWEVG-----GPGSASPTPACLLPNCCTPIFDPRVF 410
DB 283 AORLGCHTYWAVSEELLPSNGHPDGE-----VPKDEGGVFDLGGPF 325

RESULT 11
US-10-492-043-2
; Sequence 2, Application US/10492043
; Publication No. US20050019306A1
; GENERAL INFORMATION:
; APPLICANT: MOUNT SINAI SCHOOL OF MEDICINE OF NEW YORK UNIVERSITY
; TITLE OF INVENTION: A HYBRID FUSION PROTEIN TRANSCRIPTION REGULATOR TO
; FILE REFERENCE: 4225-4000PC
; CURRENT APPLICATION NUMBER: US/10/492,043
; PRIOR FILING DATE: 2004-04-05
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/352,777
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: P48-S2C protein sequence
US-10-492-043-2

Query Match      12.4%; Score 337.5; DB 17; Length 497;
Best Local Similarity 25.7%; Pred. No. 1.9e-15;
Matches 123; Conservative 71; Mismatches 185; Indels 99; Gaps 19;

QY 17 WLIGEISSCYEGLQWLDEARTCFRVPWKHFARKOLSE-ADARIFKAWAVARGRPPSR 75
DB 15 WVVEQESQFGVGVDDTAKTMFRIPWKHAGKQDPREDQDAFFKAWAIFKGY----- 69
QY 76 GGGPPPEAEATERRAGWKTFRFCALRSTRFRVMLRDNSSG--DPADPHKVVALSRELCWREGP 134
DB 70 -----KEGDTGGPAVWVKTRLCALNKSSEFKVEFPERGRMDVAEPYKYLPPGIVSGQF 124

; ORGANISM: Homo sapiens
US-10-741-600-1292

Query Match      13.3%; Score 364.5; DB 17; Length 452;
Best Local Similarity 27.4%; Pred. No. 2.2e-17;
Matches 113; Conservative 43; Mismatches 153; Indels 103; Gaps 16;

QY 10 PRVLFGWLLGEISSCYEGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVARGR 69
DB 6 PRIL--PWLVSQDLGLQLEGVAVWVKSTRFRIPWKHGLRQDAQQBDFGIFQAWAEATGA 63
QY 70 WPPSSRGSGPPPEAEATERRAGWKTFRFCALRSTRFRVMLRDNSSGDPADPHKVVALSRELC 129
DB 64 YP-----GRDKPDLPT-----WKNRPSALNKKEGLRLAEDRSKDPDHPKHYEFV----- 110
QY 130 WREGPDTQTEAEAPAAVPPPGGPPGFLAHTAGLQAPGFLPAGDGKGLLLQAVQ 189
DB 111 -NSGVG---DFSQDPTSDTNGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASWGADVPVTKAPGEGGGLPLTGACAGGPGGLPAGELYGWAVETTPSPGPQ 249
DB 137 EDILDELL-GNMVLAPLP-----DPGPP-----SLAVAPPCPQ 169
QY 250 PAALTGGERAAEPSPHQAEPYLSPPSACTAVQEP--SPGALDVTIMYKGRTVLQKVVGHP 308
DB 170 PLRSPSLDNPT-----FNLGSENPLKRLLVGGEWEFEVTAFYRGQVFOQTISCP 223
QY 309 SCTFLYGGPDPAVRATDPOQVAPPSPA--ELPDQKQLRYTEELLRHVAPGLHLELRGPOLW 367
DB 224 EGLRVG-SEVGDRTPLGWPVTLDPGMSITDRGVMSYVRHVLSCLGGGLALWRAGQWLW 282
QY 368 ARMGCKVYWEVG-----GPGSASPTPACLLPNCCTPIFDPRVF 410
DB 283 AORLGCHTYWAVSEELLPSNGHPDGE-----VPKDEGGVFDLGGPF 325

RESULT 12
US-10-276-774-1410
; Sequence 1410, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y., Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1410
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-1410

Query Match      12.3%; Score 335.5; DB 15; Length 72;
Best Local Similarity 90.1%; Pred. No. 3e-16;
Matches 64; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 152 GGPGRFLAHTAGLQAPGFLPAGDGKGLLLQAVQSCSLADHLLTASWGA-DPVPTKA 210
DB 1 GGPGRFLAHTAGLQAPGFLPAGDGKGLLLQAVQSCSLADHLLTASWGGKDPPTKA 60
QY 211 PGEQGEGLPLT 221
DB 61 LGEQGEGLPLT 71

RESULT 13
US-10-492-043-30
; Sequence 30, Application US/10492043
; Publication No. US20050019306A1
; GENERAL INFORMATION:
; APPLICANT: MOUNT SINAI SCHOOL OF MEDICINE OF NEW YORK UNIVERSITY
; TITLE OF INVENTION: A HYBRID FUSION PROTEIN TRANSCRIPTION REGULATOR TO
; FILE REFERENCE: 4225-4000PC
; CURRENT APPLICATION NUMBER: US/10/492,043
; CURRENT FILING DATE: 2004-04-05
```


Qy	122	YALSRELCWREGPGTDTQTEAEAPAAVPPQGPPGPF	LAHTHAGLQAPGLPAPAGDKGD	181
Db	66	YEFV-----NSGVG-----DFSQDTSPTDNGG-----	GSTSD	93
Qy	182	LLLQAVQQSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAWE	241	
Db	94	-----TQEDIIDELL-GNNVLAPLP-----	DPGPP-----SLA	120
Qy	242	TTSPGPGPAAUUTGAAAPESPHQAEPLYLSPSPSACTAVQBP-SPGALDVTIMYGRV	300	
Db	121	VAPEPCQPLRSPSLDNPTP-----FPNLGSPENPLKELLVPGEEWEFEVTAFYRGQV	174	
Qy	301	LQKVGHPSCTFLYGPDPDAVRATDQQAFAFSPA-ELPDQQLRYTEELLRHVAPGLHL	359	
Db	175	FOQTISCPEGLRVG-SEVGDRTLPGWPVTLFDPGMSLTDRGVMYVRHVLSCLGGLAL	233	
Qy	360	ELRGPOLWARRMKCKVYWEVG-----GPPGSASPTPACLLPRNCDTPIPDFRVF	410	
Db	234	WRAGQWLMAQRHCHTYWAVSEELLPNSGHGDGE-----	VPKDKEGGVFDLGP	284
Qy	411	F-----QELVEFRARQRGSPRYTIYLGFGQDLSAGRPKKSLLVVKLEPWLCRVHLEG	464	
Db	285	IVGSLGPPDLITFTEGSGR-SPRYALMFCVGSWPQDQWTKRLVMVKVPTCLRALVEM	343	
Qy	465	TORGVSLLDSSDLCLSSANSL-----YDDIECFMEL	499	
Db	344	ARVGGASSLENT-VDLHISNSHPLSLTSDQYKAYLQDL	380	

Search completed: May 6, 2005, 16:50:52
Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:28:18 ; Search time 41 Seconds
(without alignments)
1180.415 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731
Sequence: 1 MALAPERAAAPRVLFGEWLLG.....SANSLYDDIECFMLELEQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	780	28.6	491	2 S56753	interferon regulat
2	482.5	17.7	504	2 G02474	interferon regulat
3	440	16.1	459	2 J03520	interferon regulat
4	376	13.8	450	2 S57837	lymphoid-specific
5	333.5	12.2	424	2 A35861	interferon consens
6	328.5	12.0	393	2 A45017	transcription fact
7	306	11.2	425	2 A45064	interferon consens
8	259.5	9.5	399	2 J04592	transcription fact
9	224	8.2	203	2 S57836	lymphoid-specific
10	183	6.7	1334	2 T50568	probable multi-dom
11	176	6.4	325	2 I52998	interferon regulat
12	176	6.4	1146	2 A38587	collagen, cornea-s
13	167	6.1	349	2 A53340	interferon regulat
14	166	6.1	325	2 B31595	interferon regulat
15	166	6.1	329	2 A31595	interferon regulat
16	165.5	6.1	328	2 A36330	interferon regulat
17	160.5	5.9	1400	2 T31555	hypothetical prote
18	160	5.9	1747	2 A45974	collagen alpha 1(X
19	159.5	5.8	416	1 SKXIAG	dermal gland prote
20	159	5.8	576	2 T36729	probable serine/th
21	158.5	5.8	839	2 T04859	extensin homolog F
22	157.5	5.8	705	2 A35363	synapsin I splice
23	156.5	5.7	1464	2 S59856	collagen alpha 1(I
24	156	5.7	381	2 T27806	hypothetical prote
25	156	5.7	1857	2 S31212	collagen alpha 1(X
26	156	5.7	1888	2 S78476	collagen alpha 1(X
27	156	5.7	3149	1 QQB88	BPLF1 protein - hu
28	155	5.7	801	2 T29018	hypothetical prote
29	154	5.6	660	1 QQB83	BHLF1 protein - hu

30	154	5.6	3530	2 A59266	unconventional myo
31	150.5	5.5	691	2 A25704	synapsin I - rat
32	150	5.5	704	2 A30411	synapsin Ia - rat
33	149.5	5.5	1952	2 T48814	hypothetical prote
34	148.5	5.4	560	1 WPHUM	mullerian inhibiti
35	148	5.4	384	2 H70580	hypothetical prote
36	148	5.4	1464	1 CGHU1S	collagen alpha 1(I
37	148	5.4	1466	1 CGHU7L	collagen alpha 1(I
38	147	5.4	301	2 T21314	hypothetical prote
39	147	5.4	633	2 B40983	collagen alpha 1(X
40	146.5	5.4	1042	1 CGCH1S	collagen alpha 1(I
41	145	5.3	296	2 A31219	collagen 1 - Caeno
42	145	5.3	1207	2 T00378	XIRAA0641 protein -
43	145	5.3	2715	2 T13049	eyelid - fruit fly
44	144.5	5.3	1315	2 A56101	collagen alpha 1(X
45	144.5	5.3	1546	1 CGHU2E	collagen alpha 2(X

ALIGNMENTS

RESULT 1

S56753
interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (Chicken)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S56753
R:Grant, C.E.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995
A>Title: CIRP-3, a new member of the interferon regulatory factor (IRF) family that is r
A:Reference number: S56753; MUID:9534365; PMID:7541908
A:Accession: S56753
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <GRA>
A:Cross-references: UNIPROT:Q90643; EMBL:U20338; NID:G790580; PIDN:AAA86995.1; PID:G7905
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 28.6%; Score 780; DB 2; Length 491;
Best Local Similarity 38.5%; Pred. No. 9.8e-40;
Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY	2	ALAPERAAAPRVLFGEWLLG	ISGGYEGLOWLDEARTCFRVWKGHPARKDLSEADARIK	61
DB	3	ALDSEGDAQKLRFGPWLNAVSSGLYRGLCWDPDRIRIPRIWKGINARKDVTSSDVEIEK	62	
QY	62	AWAVAGRWPPSSRGSGPPPEAETAEACAGKTNFCALRSTRFVWLRNDSGDPADPHKV	121	
DB	63	AWAKASGRY-----EGNAEDPAKWKTNFCALRSTHFMFLLEDKNCNDDPHKV	111	
QY	122	VALSRELWCWREGPGTDTQTEAEAPAAVPPQ-----GGP-----PGPFLAHTAGLQ-----	167	
DB	112	YAVA-----SGVPNDRSGSGPVAGALQQPQLLNHDLALENTPT	152	
QY	168	-APGFLPAPAGDKGL-LQAVQQCLADHLITASWADPVPTKAPGE--QEGPLTG--	222	
DB	153	DSTEGVAAALQVDLDLLQSLVQHCNISAL-----GSQPTLWAHTGDALPEDALLPGQ	207	
QY	223	-ACAGGPGLPAGELYGVAVETTPSPGPQ-----AALTGEEAAAPESHQAE	269	
DB	208	DGCLFPG-----QFQDWRQLEPFLLLGNQPLTGGCGQDGAGALPVSECAIPAPSPAE	262	
QY	270	YL-----SPSPSACTAVQBPSPCALDVTIMYKGRITVLQKVVGHPSCFTFLYGGPDPAPVRATD	325	
DB	263	LLFQGANPAPPAGDGGFLPLLLITIIYRKVYVQEQVDDSRCLVLAQPLDPAV--AE	320	
QY	326	PQCVAFPPSPAEPLQDKQLRYTEELLRHVAPGJLHLELGRGQPLWARRMGKCKVYNEV-----	380	
DB	321	QRLVLFPSPASLPDPRRYTDLLE--VAGLRLEQRAGQLLATRLKCKKCFWALSQQLE	378	
QY	381	GGPGSASPTACLLPRNCDTPIDFRVFFQELVEFRARQGRSPRYTIYLGFGQDLSA	440	
DB	379	GGEP-----PLNLLHRDQETIIFDFRVFCTELDRDSDRSRSPDFTIFLCFGQCFSS	431	

A>Description: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon
A:Reference number: S57838
A:Accession: S57838
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-450 <MAT2>
A:Cross-references: EMBL:U11692; NID:g790199; PIDN:AAA75309.1; PID:g790200
R:Eisenbeis, C.F.; Singh, H.; Storb, U.
Genes Dev. 9, 1377-1387, 1995
A>Title: PIP, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcrip
A:Reference number: I49359; MUID:95317607; PMID:7797077
A:Accession: I49359
A>Status: translated from GB/EMBL/DBDJB
A:Molecule type: mRNA
A:Residues: 1-450 <RES>
A:Cross-references: EMBL:U34307; NID:g976446; PIDN:AAA75283.1; PID:g976447
C:Genetics:
A:Introns: 72/3; 135/1; 164/3; 213/1; 249/1; 367/1; 404/3
C:Superfamily: lymphoid-specific interferon regulatory factor
C:Keywords: alternative splicing

Query Match 13.8%; Score 376; DB 2; Length 450;
Best Local Similarity 28.1%; Pred No. 1.5e-15;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

Qy 16 EWLLGEISSGCGYGLQMLDEARTCFRPVWKHFARKDLS-EADARIFKAWAVARGWRPSS 74
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
26 QWLIDQIDSGKYPLGWENEEKSVFRIPWKHAGKDQYNREEDAALFKAWLFKGKF--- 81
Qy 75 RGGPPPEAEATERAGWKTNRFCALRSTRFVMRLDNSG-DPADPHKVYALSRELCHWEG 133
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
82 REGIDKKDPPT-----WKTRLRCALKNSDFEELVERSQLDISDPYKVYRI----- 127
Qy 134 PGTDQTAEAPAAVPVPGGPPGP----FLATHAGLAQGAPLPAPAGDKGLLIQAQQQS 190
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
128 -----VPEGAKKGAKQLTLDLTQMAMGHYPWTAP----YGSLPAQQVHNY 169
Qy 191 CLADHLITASMGADPVPTKAPGEGQGLTLTGACAG----GPGLP-----AGELYGWAVE 241
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
170 MMPFH--DRSW-RDYAPDSHPPIPYQCFTVGPRHHWQGCSCENGCVGTGFYACAPP 226
Qy 242 TTPSPG-PQPALTTGEEAAEPESHQABPYLSPSPSACTAVQEPSGALDVNTMYKGRVT 300
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
227 ESQAFGPIEPSIRSAEAAL-----LSDC-----RLHTCLYYRDILV 263
Qy 301 LOKVVUGH-PCTFLYGGPDPAVRATDPOOVAFPSPAELPQOKOLRYTEELLRHVAPGLHL 359
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
264 KELTTTSPEGCRISHG---HYTVSNLDQVLFPYP---DDNGQRKNIEKLSHLERGLVL 317
Qy 360 ELRGPOLWARMRGMCKVYVEGVGGPPGSAS--PSTPACILLPNRCTPIFDPRVFPELVFEF 417
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
318 WMAPDGLYAKLCOSRIYWD--GPLALCSDRPNK-----LEDOTCKLFDTCQFISELQVF 371
Qy 418 RARQRGRSPRTIYLGFQDLISAGRKPESKSLVLVKLEPWLCR 459
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
372 -AHHRGPARPRFQVTLCPFGEFFPDQ-RORKLITAHVEPLLAR 411

RESULT 5
A35861
interferon consensus sequence-binding protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 09-Jul-2004
C:Accession: A35861
R:Driggers, P.H.; Emnist, D.B.; Gleason, S.L.; Mak, W.H.; Marks, M.S.; Levi, B.Z.; Flanc
Proc. Natl. Acad. Sci. U.S.A. 87, 3743-3747, 1990
A>Title: An interferon gamma-regulated protein that binds the interferon-inducible enhan
A:Reference number: A35861; MUID:90251633; PMID:2111015
A:Accession: A35861
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-424 <ARI>
A:Cross-references: UNIPROT:P23611; GB:M32489; NID:g194088; PIDN:AAA337878.1; PID:g330932

```
Qy 76 GGGPPPEAETAAERAGWKTNRFCALRSTRFVMLRDNSSG-DPADPHKYVALSRELCEWREGP 134
Db 70 -----KEGDTGGPAVWKTRLCALNKSSEFKEVPERGEMDVABEYKYQLLPGLVSGQP 124
Qy 135 GTDQTEAABAAVPPPOGGPPGLFATHAGLQAPGLPAPAGDKGLLLOAQQOOSCLAD 194
Db 125 GTQKVPKSG-----RQHSSVS-----SERKEED-----AMQNCTLSP 156
Qy 195 HLLTASGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGPQPAALT 254
Db 157 SVLODSLNEE-----EGASG-----GAVHSDIG-----SSSSSSPPEQVETDT 196
Qy 255 TGEAAAPESPQAB-----PYLSPSPSACTAQVPSPG-ALDVTIMYKGRVLOKVGHP 309
Db 197 T-----EAPFGQDORSLEFLP-----PEPYSILLTFIYNGRVVGAQVOSLD 240
Qy 310 CTFLYGPPDPAVRATDPOQVAFPSPAELPQKQQLRYTEILLRHVAPGLHLELRGPQLWAR 369
Db 241 CRLV---AEPSSGSSMEQVLPKPGPLEP-----TORLLSQLERGILVASNPRGLFVQ 291
Qy 370 RMGCKKYWEVGGPPGSASTPACILPRNCDTPIFDFRVFFQELVFEFRARORRG-SPRY 428
Db 292 RLCPIPSWNAPOAPPQGGPH-----LLPSNECEVELFETAYFCRDLV--RYFQGLGPPPKF 345
Qy 429 TIVLFGQDLSAGRPKEKSLVVKLEPWLCKRVHLEGT--OREGVSSL 473
Db 346 QVTLNFWEEBSHGSHHTQNLITVMEQAFARYLLEQTPEQQAILLS 392

RESULT 7
A45064
interferon consensus sequence binding protein, ICSBP (DNA-binding domain) - human
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998
R;Weisz, A.; Marx, P.; Sharf, R.; Appella, E.; Driggers, P.H.; Ozato, K.; Levi, B.Z.
J. Biol. Chem. 267, 25589-25596, 1992
A;Title: Human interferon consensus sequence binding protein is a negative regulator of
A;Reference number: A45064; MUID:93094284; PMID:1460054
A;Accession: A45064
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-425 <WEI>
A;Experimental source: lung and blood
A;Note: sequence extracted from NCBI backbone (NCBI:P120312)
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription regulation

Query Match 11.2%; Score 306; DB 2; Length 425;
Best Local Similarity 23.4%; Pred. No. 2.3e-11;
Matches 121; Conservative 67; Mismatches 174; Indels 156; Gaps 21;

Qy 16 EWLGEISSCYEGLOWLDEARTCFRVPWKHFARKDLSE-BADARIFKAWAVARGWRPSS 74
Db 12 QWLIEQIDSSMYPGLIWENEESKMFRIPWKHAGQDYNEVDASIFKAWAVFKGK---- 67
Qy 75 RGGPPPEAETAAERAGWKTNRFCALRSTRFVMLRDNSSG-DPADPHKYVALSREL----- 128
Db 68 -----KEGDKAEPATWKTRLCALNKSDFEEVTDRLSDISEPYKYRIVPEEQKCK 121
Qy 129 -----CWREGPQDQTEAABAAVPPPOGGPPGLFATHAGLQAPGLPAPAGDKGD 181
Db 122 LGVATACVNEVTEMECGRESEIDELIKEPS-----VDDYMGIMKRSPPDA----- 168
Qy 182 LLLQAVOOSCLADHLLTASGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVE 241
Db 169 -----CRSQ--LLPDW-----WAHE 181
Qy 242 TTSPGPQPAALTGEEAAAPESPQABEYLYSPSPSACTAVQEPSGALDVTIMYKGRV 301
Db 182 --PSTGRR--LVTGYTTY-DAAHSAP-----SQWISFYVYGGKLVG 217
```

```
Qy 302 QKVVGHP-SC-----TFLYGPPDPAVRATDPOQVAFPSPAELPQKQQLRYTEEL 349
Db 218 QATTTCPEGCRLSQSGPLPGTKLYGP-----EGLELVRFPPADTIPSERQROVTRKL 270
Qy 350 LRIVAPG-LHLEURGOLWARRMGKCKYVWEVGGPPGSGASPSPTACILPRNCDTPIFDPR 408
Db 271 FGHLERGVLHSSRQG--VFVKRLCQGRVFCSGNAVVCGRPNK-----LERDEVVQVFDTS 324
Qy 409 VFQELVEFRARORRGSPRYTIVLFGQDLSAGRPKEKSLVVKLEPWLCKRVHLEGTQRE 468
Db 325 QFFRELQQFYNSQGR-LPDGRVVLCEGEEFPDMPALRSKLLVQIE-----QLYVRQLAEE 379
Qy 469 GVSSDSSDL-----DLCLSSANSLY 489
Db 380 AGKSCGAGSVMQAPBEPFPDQVFRMPFDICASHQRSFF 417

RESULT 8
JC4592
transcription factor ISGF3 gamma chain - mouse
N;Alternate names: interferon-stimulated gene factor 3 gamma chain
C;Species: Mus musculus (house mouse)
C;Date: 10-Apr-1996 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
C;Accession: JC4592; S71599
R;Suhara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.
J. Biochem. 119, 231-234, 1996
A;Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma/p48)
A;Reference number: JC4592; MUID:97037063; PMID:8882710
A;Accession: JC4592
A;Molecule type: mRNA
A;Residues: 1-399 <SUH>
A;Cross-references: UNIPROT:Q61179; EMBL:U51992; NID:g1263309; PIDN:AAC52494.1; PID:g1361
A;Experimental source: L929 cells
R;Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, M.
FEBS Lett. 358, 225-229, 1995
A;Title: Possible involvement of the transcription factor ISGF3gamma in virus-induced ex
A;Reference number: S71599; MUID:95145714; PMID:7843405
A;Accession: S71599
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-87 <KAW>
C;Genetics:
A;Gene: Isgf3g
A;Map position: 14
C;Function:
A;Description: responsible for specific interaction with the promoter element, interferon
ed gene factor 3, the primary regulator of type I interferon responses; involved in type
A;Note: induced by interferon-alpha and interferon-beta
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; signal transduction; transcription factor
P;10-117/Domain: DNA binding #status predicted <DNB>

Query Match 9.5%; Score 259.5; DB 2; Length 399;
Best Local Similarity 22.5%; Pred. No. 1.3e-08;
Matches 105; Conservative 64; Mismatches 192; Indels 105; Gaps 14;

Qy 17 WLLGEISSCYEGLOWLDEARTCFRVPWKHFARKDLSE-ADARIFKAWAVARGWRPSSR 75
Db 15 WIVEQESGHFGPVCVDDAAKTNFRIPWKHAGQDFREDQDAAFKAWALFKEK----- 68
Qy 76 GGGPPPEAETAAERAGWKTNRFCALRSTRFVMLRDNSSG-DPADPHKYV-----ALSREL 128
Db 69 -----HKDGDIGHPAVWKTRLCALNKSSEFEEVPERGRMDVABEYKYRILPAGTLNPQ 124
Qy 129 CWREGPQDQTEAABAAVPPPOGGPPGLFATHAGLQAPGLPAPAGDKGLLLOAVQ 188
Db 125 RNQSPCKRSISCVSPERBENMENGRTNGVNHSDSGSNIGG-----GGNGS----- 171
Qy 189 QSLADHLLTASGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGP 248
Db 172 -----NRSDSNCSNCELEEG-----AG----- 189
Qy 249 QPAALTGEEAAAPESPQABEYLYSPSPSACTAVQEPSGALDVTIMYKGRVTLQKVVGHP 308
```

```
Db 190 -----TTEATIREDFVLEHOLPLN-----SDYSLLTFTYGGRVVGKTVHSL 233
Qy 309 SCTFLYGGPDPAVRATDPOQVAFPPGPAELPDQKQLRYTEELRHVAPGLHLELPGPQLWA 368
Db 234 DCRUVAERSD---SESSMEQVEFPK-----DPLEPTQHLLNQLDRGVLVASNSRGLEFV 284
Qy 369 RMGKCKVYVGGPPGSPGASPTACLLPRNCDTPIFRVFBELVFRARQRRG-SPR 427
Db 285 ORLCPIPIISWNAPEAPPGPGRH---LLPSNKCVELFKTTFYCRDLAQV---FQGGPPPK 338
Qy 428 YTIYLGQODLSAGRPKESLVLVKLEPWLCHVHLEGTQREGVSSL 473
Db 339 FOATHLFEESPGSSHQENLITVQMEQAFARHLEKTIPEEKAAL 384

RESULT 9
S57836
Lymphoid-specific interferon regulator factor - mouse (fragments)
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S57836
R:Matsumura, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Kiefer, F.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A:Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A:Reference number: S57836; MUID:95334364; PMID:7541907
A:Accession: S57836
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-203 <MAT>
A:Cross-references: EMBL:U11692
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 8.2%; Score 224; DB 2: Length 203;
Best Local Similarity 34.8%; Pred. No. 8.7e-07;
Matches 57; Conservative 22; Mismatches 49; Indels 36; Gaps 6;

Qy 16 EWLGEISSCYEGQLWLDARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPSS 74
Db 4 QMLIDQSGKYPGLWENEEKSVFRIPWKHAGKQDYNEEDALFKAWLFKGF---- 59
Qy 75 RGGGPPPEAETAEAGKTNFRCALSTRFRVWLRDMSG-DPADPHKYVALSRELCWREG 133
Db 60 REGIDKPDPT-----WKLRLCALNKSNDFFELVERSQLDISDPYKVYRI----- 105
Qy 134 PGTDQTEAEAPAAVPPQGGPGPFLAHTHAGI---QAPGLPLA 174
Db 106 -----VPEGQRKNIEKLLSHLERGLVLMMPDGLYA 136

RESULT 10
T50568
Probable multi-domain regulatory protein [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50568
R:Redenbach, M.; Kieser, H.M.; Denapaita, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: T50568
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1334 <RED>
A:Cross-references: UNIPROT:Q9RKB9; EMBL:AL133220; PIDN:CAB61705.1
A:Experimental source: strain A3(2)
C:Genetics:
A:Note: SCC75A.05c

Query Match 6.7%; Score 183; DB 2: Length 1334;
Best Local Similarity 28.1%; Pred. No. 0.0018;
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;
```

```
Qy 18 LLGEISSCYEGQLWLDARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPSSRGG 77
Db 146 VLADLDRTAEAARW---ETRHFEALRARHTAALDGOAEHSLPELTALCDGH----- 195
Qy 78 GPPPEAETAEAR-----AGWKNFRCALSTRFRVWLRDMSG-DPA-----D 117
Db 196 -PLDEPQALRLRALRDSGRTAELAAAYEAVRR--LLADRLGTDPGPELRLHAELLSPS 252
Qy 118 PHKYVALSRELCWREGP-----CTDQTEAEAPAAVPPQGGP-PGPFLAHTHAGL 166
Db 253 FTTPTGRSRTPGWTSPPGPGASGAGAAAGCTDVASGAGAAAGDDPASGASPAVAPGSGGG 312
Qy 167 QAPGLPAPAGDKGDLLOAVQOCLADHLITASWGADPVPTKABCEQEGLEPLTGACAG 226
Db 313 PAPGWPAPGTAPGSSSTAPPHTTAAAD---TA-----PAPGPTSAPGTA---PAAGTAAP 362
Qy 227 GPGL--PA-GELY--GWA--VETTPSPGPQPAALTTGEA-----AAPESPHQAEPYL 271
Db 363 APGTAGPAPGTSYAPGTAPVAGTTTAPGTAPGTAPGTAGTAPRTSYAPGTAPVAGTTTAPGT 422
Qy 272 SPSPSACTAV-----QEPSFGALDVTIMYKGRVTLQKVVGHPSCCTFLYGGPPDPVAVRAT 324
Db 423 APAPGSTPAGTVPAGTAPAGPQPA---DGR---RPVTGPASGT---GPG---AAT 468
Qy 325 DPQQVA-----FPSPA 335
Db 469 PPEAAAAAAGSAPSPA 485

RESULT 11
I52998
Interferon regulatory factor 1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I52998
R:Cha, Y.; Sims, S.H.; Romine, M.F.; Kaufmann, M.; Deisseroth, A.B.
DNA Cell Biol. 11, 605-611, 1992
A:Title: Human interferon regulatory factor 1: intron/exon organization.
A:Reference number: I52998; MUID:93000481; PMID:1382447
A:Accession: I52998
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: UNIPROT:P10914; GB:L05072; NID:G184648; PIDN:AAA36043.1; PID:G184649
C:Genetics:
A:Gene: GDB:IRF1
A:Cross-references: GDB:127269; OMIM:147575
A:Map position: 5q31.1-5q31.1
A:Introns: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1

Query Match 6.4%; Score 176; DB 2: Length 325;
Best Local Similarity 23.6%; Pred. No. 0.0011;
Matches 69; Conservative 35; Mismatches 108; Indels 80; Gaps 12;

Qy 11 RVLFGEWLGEISSCYEGLOWLDEARTCFRVPWKHFARK--DLSEADARIFKAWAVARG 68
Db 5 RMRMRPWLQEMQINSQIPLGIWINKKEIMFQIPWKHAAKHGWDINK-DACLERSWAIHTG 63
Qy 69 RWPPSSRGGPPPEAETAEAGKTNFRCALSTRFRVWLRD---NSGDDPADPHKYVALS 125
Db 64 RY-----KAGEKEPDPT-----WKNFRCAAMNSLPDIEVKQDSRNKGSAA--VRVIRML 112
Qy 126 RELC---WREGPGTDQTEAEAPAAVPPQGGPPGPF-----LAHTHAGLQAPGLPLA 174
Db 113 PPLTNQKQKRSKSRDAKSKAKKSCGDSPTDFSDGLSSSTLPDDHSSYTVFG----- 168
Qy 175 PAGDKGDLLOAVQOCLADHLITASWGADPVPTKAPGEGQGLPLTGACAGGGLP---- 231
Db 169 -----YMQDLEVEQALT-----PALSPCAVSSTLPDWH 196
Qy 232 -----AGELYGWAVETTPSPGPQPAALTTGEAAAPESPHQAEPYLSPS 274
Db 197 IPVEVVPDSTSLYNFQV-----SPMPTSEATTDEEGKLPEDINKLLEQS 244
```

RESULT 12

A38587
collagen, cornea-specific - chicken
C;Species: Gallus gallus (chicken)
C;Date: 14-Feb-1992 #sequence_revision 15-Aug-1997 #text_change 09-Jul-2004
C;Accession: S16501; A38587
R;Marchant, J.K.; Linsemayer, T.F.; Gordon, M.K.
Proc. Natl. Acad. Sci. U.S.A. 88, 1560-1564, 1991
A;Title: cDNA analysis predicts a cornea-specific collagen.
A;Reference number: A38587; MUID:91142213; PMID:1705041
A;Accession: S16501
A;Molecule type: mRNA
A;Residues: 1-1145 <NAR>
A;Cross-references: UNIPROT:Q90584; EMBL:M60172; NID:9211609; PIDN:AAA48703.1; PID:92116
A;Accession: A38587
A;Molecule type: mRNA
A;Residues: 1-174, 'X', 176-233, 'X', 235-344, 'X', 346-408, 'X', 410-499, 'X', 501-876, 'X', 878-1119
A;Cross-references: GB:M60172
C;Keywords: cornea

Query Match 6.4%; Score 176; DB 2; Length 1146;
Best Local Similarity 23.8%; Pred. No. 0.004;
Matches 109; Conservative 31; Mismatches 132; Indels 186; Gaps 26;
QY 66 ARGWRPSSRG-----GPPPEAETABRAGKTNFRCALRSTRFRVLMRDNSGDPAD 117
DB 348 ARGPPGSGDTGEPGLTGPPGQPLGPNCRPGAK-----GEPGA 387
QY 118 PHKYALSLRCLWREGTDTQTRAEAPAVPPGPGPFLAHTAGLQAP-GP--LP- 173
DB 388 PGKVIS-----AEGSSTIALPGPP-GPPGPIGTGPPGVPVGPAGLPG 431
QY 174 --APADKG--DILLQAV--QSCADHLLTASWG--ADPVTKAPGEQEGPLTGAC 224
DB 432 QQGRGSGKSAVEVVIETIKTEVSSLASQMLSDQGRAGPPGPPGSGVQGLP---GP 488
QY 225 AGGPGPLPAGLYGWAVETTPS--PGQPAALTGCE--AAAPESP-----HQAEPYLS 272
DB 489 RGPPLPG-----PSGPPRGSSVSSEITVSGPPGPPGPKGQGE----- 533
QY 273 PPSACTAVQBPSPGALDVTIMYKGRVLQKVGHPSCTFLYGPDPVAVATDP----- 326
DB 534 PGRGFTG--EPGEPGLPGSSHGTVTMOQPP-----GPPGPKDAGVGPAGPIP 585
QY 327 -----QQV-----AFTSPALPDOKQLRYTEELL-----RHVAPGLHLELRGP 364
DB 586 GTSRGSRIQGPFGPPGPPGPGGSSQEIQQVYADYKSDNVHRHLYTG----- 637
QY 365 QLWARRMGKCKVYVEVGGPPGSGASPTPACLLPNCIDTPIFD----- 407
DB 638 -----VQSPFG--PPGPPGIL--TTADCKNFDFAELATRVMSYVTSDDH 678
QY 408 -----RVFQELVFRARQ-----RGSP 426
DB 679 YQSPASSVSTTSVLYQELNMLQREIRQYLVGRGPP 716

RESULT 13

A53340
interferon regulatory factor 2 - human
N;Alternate names: transcription repressor IRF2
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A53340; S06894; A32828
R;Chia, Y.; Deisseroth, A.B.
J. Biol. Chem. 269, 5279-5287, 1994
A;Title: Human interferon regulatory factor 2 gene. Intron-exon organization and function.
A;Reference number: A53340; MUID:94148994; PMID:8106512
A;Accession: A53340
A;Status: not compared with conceptual translation
A;Molecule type: DNA

A;Residues: 1-349 <CHA>

A;Cross-references: UNIPROT:PI4316; GB:L24442
R;Itoh, S.; Harada, H.; Fujita, T.; Mimura, T.; Taniguchi, T.
Nucleic Acids Res. 17, 8372, 1989

A;Title: Sequence of a cDNA coding for human IRF-2

A;Reference number: S06894; MUID:90045964; PMID:2813069

A;Accession: S06894

A;Molecule type: mRNA

A;Residues: 1-57, 'R', 59-349 <ITO>
A;Cross-references: EMBL:X15949; NID:G33966; PIDN:CAA34073.1; PID:G33967

R;Harada, H.; Fujita, T.; Miyamoto, M.; Kimura, Y.; Maruyama, M.; Furia, A.; Miyata, T.;

Cell 58, 729-739, 1989

A;Title: Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to

A;Reference number: A32828; MUID:89354547; PMID:2475256

A;Accession: A32828

A;Molecule type: mRNA

A;Residues: 1-68, 'I', 70-96, 'R', 98-130, 'ER', 133-152, 'GF', 155-163, 'A', 165-188, 'D', 190-210, 'A',

8, 'T', 310-313, 'PAPV', 318-319, 'TP', 322-349 <HAR>

A;Cross-references: GB:J03168; NID:G198455; PIDN:AAA9333.1; PID:G993676

A;Experimental source: clones 2 and 5

C;Genetics:

A;Gene: GDB:IRF2

A;Cross-references: GDB:127270; OMIM:147576

A;Map position: 4q35.1-4q35.1

C;Keywords: DNA binding; transcription regulation

Query Match 6.1%; Score 167; DB 2; Length 349;

Best Local Similarity 35.0%; Pred. No. 0.004;

Matches 36; Conservative 18; Mismatches 39; Indels 10; Gaps 3;

QY 11 RVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPKHFKARDLS-EADARIFKAWAVARGR 69

DB 5 RMRMPWLEEQINSNTIPGLKWLNEKKTFQIPWMAARHGWDVKDAPLFRNWAHTGK 64

QY 70 WPSSRGSGPPPEAETABRAGKTNFRCALRSTRFRVLMRDNS 112

DB 65 HQP-----GVDPKDPKT-----WKNFRCAWNSLPDIEEVKDKS 98

RESULT 14

B31595

interferon regulatory factor 1 - human

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004

C;Accession: B31595; S04075

R;Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;

Cell 54, 903-913, 1988

A;Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifically

A;Reference number: A90903; MUID:88311092; PMID:3409321

A;Accession: B31595

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-325 <M12>

A;Cross-references: UNIPROT:PI0914

R;Maruyama, M.; Fujita, T.; Taniguchi, T.

Nucleic Acids Res. 17, 3292, 1989

A;Title: Sequence of a cDNA coding for human IRF-1.

A;Reference number: S04075; MUID:89263736; PMID:2726461

A;Accession: S04075

A;Molecule type: mRNA

A;Residues: 1-325 <MAR>

A;Cross-references: EMBL:X14454

C;Keywords: DNA binding; transcription regulation

Query Match 6.1%; Score 166; DB 2; Length 325;

Best Local Similarity 23.4%; Pred. No. 0.0043;

Matches 67; Conservative 33; Mismatches 106; Indels 80; Gaps 12;

QY 17 WLIGETSSGCGEGLQWLDEARTCFRVPKHFKARK--DLSEADARIFKAWAVARGRPPSS 74

DB 11 WLEMQINSQIPGLIWINKEEMILEIPWMAARHGWDVKDAPLFRNWAHTGK 65

QY 75 RGGGPPPEAETABRAGKTNFRCALRSTRFRVLMRD---NSGDPADPHKYALSLRELCL 129


```

Db      66 KAGEKEPDPKT-----WKANFRCAMNSLPDIEVKDQSRNKGSSA--VRVYRMLPPLTKN 118
QY      130 -WREGPGTQTEAEAPAAVPPQGGPPGF-----LAHTHAGLQADGPLPAPAGDKG 180
Db      119 QRKERSKSRDAKSKAKKSCGDSFDTFSDGLSSSTLPDDHSSYTFG-----168
QY      181 DLLLQAVQOQSLADHLLTASWADPVPTKAPGEGQGLPLTGACAGGPGLP-----231
Db      169 -----YMQDLEVEQALT-----PALSPCA VSTLPDWHIPVEV 202
QY      232 ---AGELYCWAVETTFPGPQPAALTGHAAPESPHQAEPLYSPS 274
Db      203 PDSTSLYNFQV-----SPMPSISEATDEDECKLPEDIMKLEQS 244

RESULT 15
A31595
interferon regulatory factor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:Accession: A31595
R:Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;
Cell 54, 903-913, 1988
A:Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifies
A:Reference number: A90903; MUID:88311092; PMID:3409321
A:Accession: A31595
A:Molecule type: mRNA
A:Residues: 1-329 <MY>
A:Cross-references: UNIPROT:P15314; GB:M21065; NID:g198458; PIDN:AAA39334.1; PID:g293677
C:Keywords: DNA binding; transcription regulation

Query Match      6.1%; Score 166; DB 2; Length 329;
Best Local Similarity 21.6%; Pred. No. 0.0043;
Matches 65; Conservative 40; Mismatches 88; Indels 108; Gaps 12;

QY 11 RVLFGWLLGEISSGCGYGLQWLDEARTCFRVPWKHFARK--DLSEADARIFKANAVARG 68
Db 5 RMRMRPWLEMQINSQIIPGLIWKKEEMIFQIPWKHAAGWDINK-DACLFRSWAIHTG 63
QY 69 RWPPSSRGGPPPEAETAEAGWKTNFRCALRSTRFVMLRD---NSGDPADPHKYVAL- 124
Db 64 RY-----KAGEKEPDPKT-----WKANFRCAMNSLPDIEVKDQSRNKGSSA--VRVYRML 112
QY 125 -----SRELWCWREGPGT-----136
Db 113 PPLTRNQRKERSKSRDTKSKTKRKLCDGVSPDTFSDGLSSSTLPDDHSSYTTQGYLQ 172
QY 137 -----DQTEAEAPAAVPPQGGPPGFPLAHTAGLQ-----APGPLPAPA 176
Db 173 DLDMERDITPALSPCVV-----SSLSSEWHMQMDIIPDSTTDLYNLQVSPMPSTSEA 224
QY 177 GDKGDLLOAVQOQSLADHLLTASWADPVPTKAPGEGQGLPLTGACAGGPGLPAGELY 236
Db 225 ATDED-----EGKTAEDLMKLFQSEWQPTHIDGKY-----LLNEFGTQLSSVY 270
QY 237 G 237
Db 271 G 271
```

Search completed: May 6, 2005, 16:45:40
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:42:52 ; Search time 15 Seconds
(without alignments)
1390.840 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSLYDDIECFLEMELEQPA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2715	99.4	503	1	IRF7_HUMAN	Q92985 homo sapien
2	1562	57.2	457	1	IRF7_MOUSE	P70434 mus musculus
3	780	28.6	491	1	IRF3_CHICK	Q90643 gallus gall
4	487.5	17.9	498	1	IRF5_HUMAN	Q13568 homo sapien
5	458	16.8	427	1	IRF3_HUMAN	Q14653 homo sapien
6	444.5	16.3	497	1	IRF5_MOUSE	P56477 mus musculus
7	438	16.0	419	1	IRF3_MOUSE	P70671 mus musculus
8	436	16.0	467	1	IRF6_MOUSE	P97431 mus musculus
9	431	15.8	467	1	IRF6_HUMAN	O14896 homo sapien
10	376	13.8	450	1	IRF4_HUMAN	Q64287 mus musculus
11	363	13.3	451	1	IRF4_MOUSE	Q15306 homo sapien
12	338	12.4	425	1	ICSB_CHICK	Q90871 gallus gall
13	333.5	12.2	424	1	ICSB_MOUSE	P23611 mus musculus
14	328.5	12.0	393	1	IRIF_HUMAN	Q00978 homo sapien
15	308.5	11.3	426	1	ICSB_HUMAN	Q02556 homo sapien
16	259.5	9.5	399	1	IRTF_MOUSE	Q61179 mus musculus
17	176	6.4	325	1	IRF1_HUMAN	P10914 homo sapien
18	167	6.1	349	1	IRF2_HUMAN	P14316 homo sapien
19	167	6.1	1083	1	T2D3_HUMAN	O00268 homo sapien
20	166	6.1	329	1	IRF1_MOUSE	P15314 mus musculus
21	166	6.1	349	1	IRF2_MOUSE	P23906 mus musculus
22	165.5	6.1	328	1	IRF1_RAT	P23570 rattus norv
23	164	6.0	348	1	IRF2_CHICK	Q98925 gallus gall
24	163	6.0	313	1	IRF1_CHICK	Q90876 gallus gall
25	159.5	5.8	439	1	XP2_XENLA	P17437 xenopus lae
26	157.5	5.8	705	1	SYN1_HUMAN	P17600 homo sapien
27	156	5.7	1888	1	CAIE_CHICK	P03186 epstein-bar
28	156	5.7	3149	1	TEGU_EBV	P03186 epstein-bar
29	155	5.7	670	1	SYN1_MOUSE	O88935 mus musculus
30	154	5.6	560	1	YH11_EBV	P03181 epstein-bar
31	154	5.6	3530	1	MY15_HUMAN	Q9ukn7 homo sapien
32	152.5	5.6	1464	1	CA13_MOUSE	P08121 mus musculus
33	150	5.5	704	1	SYN1_RAT	P09951 rattus norv

ALIGNMENTS

RESULT 1

IRF7_HUMAN	STANDARD; PRT; 503 AA.
ID	Q92985; O00331; O00332; O00333; O75924;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Interferon regulatory factor 7 (IRF-7).
GN	IRF7.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]_TaxID=9606;
RP	SEQUENCE FROM N.A. (ISOFORM A).
RC	TISSUE=Spleen;
RA	Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,
RA	Sutherland G.R., Mak T.W.;
RL	Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX	MEDLINE=97459673; PubMed=9315633;
RA	Zhang L., Pagano J.S.;
RT	"IRF-7, a new interferon regulatory factor associated with Epstein-
RT	Barr virus latency".
RL	Mol. Cell. Biol. 17:5748-5757(1997).
RN	[3]
RP	SEQUENCE FROM N.A. (ISOFORM D).
RX	MEDLINE=99003279; PubMed=9786932;
RA	Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pitha P.M.;
RT	"Characterization of the interferon regulatory factor-7 and its
RT	potential role in the transcription activation of interferon A
RT	genes";
RL	J. Biol. Chem. 273:29210-29217(1998).
CC	-!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
CC	STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q
CC	PROMOTER (QP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1).
CC	-!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC	-!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B/BETA, C/GAMMA
CC	AND D/H; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SPLEEN, THYMUS, AND
CC	PERIPHERAL BLOOD LEUKOCYTES.
CC	-!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; U73036; AAB17190.1; -
DR	EMBL; U53830; AAB80686.1; -
DR	EMBL; U53831; AAB80688.1; -
DR	EMBL; U53832; AAB80690.1; -


```

QY 479 DCLSSANSYDDIECFLEMLQ 501
DB 433 GLCLSSNSLYEDIEHFLMDLQ 455

RESULT 3
IRF3_CHICK
ID IRF3_CHICK STANDARD; PRT; 491 AA.
AC Q90643;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Interferon regulatory factor 3 (IRF-3).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95334365; PubMed=7541908;
RA Grant C.E., Vasa M.Z., Deeley R.G.;
RT "cIRF-3, a new member of the interferon regulatory factor (IRF)
RL Nucleic Acids Res. 23:2137-2146(1995).
CC -!- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U20338; AAA86995.1; -
DR HSSP; P15314; 11F1.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF, 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF, 1.
DR SMART; SM00348; IRF, 1.
DR PROSITE; PS00601; IRF, 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator.
FT DNA BIND 14 112 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 491 AA; 54441 MW; 54441 MW; CAEC2AA8BE976D9 CRC64;

Query Match 28.6%; Score 780; DB 1; Length 491;
Best Local Similarity 38.5%; Pred. No. 6.6e-38;
Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY 2 ALAPERAARVLEWLGIEISSGVEGLQWLDEARTCFRVPWKHFARKDLSADARIFK 61
DB 3 ALDSEGAOKLRRGPNLLNAVSSGLYRGLCWIDPDRRIFRIPWKHNRKDVTSDDVEIFK 62
QY 62 ANAVARGRPSPSGGPPPEAETARAGWKTFRCALRSTRFVLMRDSNGDPADPHKV 121
DB 63 AWAKASGRY-----EGNAEDPAKWTNFRCALRSTHFMFLERDKCNDDPHKV 111
QY 122 YALSRELWCWREGPTDQTEAEAPAAVPPQ--GGP-----PGPFLATHAGLQ---- 167
DB 112 YAVA-----SGVPNDRGSGVPVAGALQQQPQLLNHHDLALENTPT 152
QY 168 -APGPIAPAGDKGLD-LIQAVQOSCLADHLLTASWAGDPVPTKAPGEG--QGLPLTG- 222
DB 153 DSTEGVAAAALTVQDLQLQSLQHCNISAL-----GSQPTLWAHTGDALPEDALLPGQ 207

```

```

QY 223 -ACAGGPGLPAGELYG-AVETTPSPGPQP-----AALTGGAAPESPHQARP 269
DB 208 DGCLPGP-----QFQDMRQLBEPILLGNQPLTGGCGGQDAGALPVSECAIPAPSPARE 262
QY 270 YL-----SPSPSACTAVQPSGALDVTIMYKGRVTLVQKVGHPSCTFLYGGPDPDPAVATD 325
DB 263 LLFQSANPAPPPAGDIGGLPLLDITIIYRGKMYQSQVDDSRCLVAYQPLDPAV--AE 320
QY 326 PQQVAFPSPAELPDQKLYRTEELRHVAPGLHLRGPQLWARRMGKCKVYEV----- 380
DB 321 QRLVLFPSPASLPDRQRRTYTEDLLE--VAGLRLFSRAGQLLATRLKKCKVFWALSQQLE 378
QY 381 GGPASPSPTACLLPRNCOTPIPDFRVFQELVEFRARQRGSPRTIYLGFGQDLQA 440
DB 379 GGEP-----PLNLLHRDQETTIFDFRVCTELDRFRDRSRERSPDFTIFLCGQCFSS 431
QY 441 GRPKESLVLVKLEPWLRCRVHLEGTQREGVSLDSDLDLCLSSANSYDDIECFLEMLE 500
DB 432 TKPKESKLLVLKLVFPQFCEYWEYQVQGGASSNGVNSQLSDSFNLFELIEQYHMQTD 491

RESULT 4
IRF5_HUMAN
ID IRF5_HUMAN STANDARD; PRT; 498 AA.
AC Q13568; Q9BQF0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN IRF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grossman A., Mittrucker H.W., Lantonio L., Mak T.W.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51127; AAA96056.1; -
DR EMBL; BC004201; AAH04201.1; -
DR EMBL; BC004139; AAH04139.1; -
DR HSSP; P23906; 11RF.
DR Genew; HGNC:6120; IRF5.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF, 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF, 1.
DR SMART; SM00348; IRF, 1.
DR PROSITE; PS00601; IRF, 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 16 118 TRYPTOPHAN PENTAD REPEAT.
FT DOMAIN 142 149 POLY-GLU.
FT CONFLICT 161 175 EDVKWPTLQPTLR -> DAVQSGPHMTPTYSLLKEDVKW
SQ SEQUENCE 498 AA; 56044 MW; 01B2ED95C28384E8 CRC64;

```



```
QY 479 DLCLSSANSL---YDIECFMLMEL 499
DB 392 DLHISNHPLSLTSQYKAYLQDL 415

RESULT 6
IRF3_MOUSE
ID IRF3_MOUSE STANDARD; PRT; 497 AA.
AC P56477;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN IRF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lymph node;
RA Grossman A., Kondo S., Antonio L., Mak T.W.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC 1- SUBCELLULAR LOCATION: Nuclear.
CC 2- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF028725; AAB81997.1; -
CC HSSP; P23906; IIRF.
CC MGD; MGI:1350924; Irfs.
CC InterPro; IPR001346; IRF.
CC Pfam; PF00605; IRF; 1.
CC PRINTS; PR00267; INTERNEGCT.
CC ProDom; PD002355; IRF; 1.
CC SMART; SM00348; IRF; 1.
CC PROSITE; PS00601; IRF; 1.
CC Transcription regulation; DNA-binding; Nuclear protein.
KW TRANSCRIPTION PENTAD REPEAT.
FT DOMAIN 9 12 POLY-PRO.
FT DOMAIN 140 148 POLY-GLU.
SQ SEQUENCE 497 AA; 56005 MW; D8BD54DB946E354F CRC64;

Query Match 16.3%; Score 444.5; DB 1; Length 497;
Best Local Similarity 29.2%; Pred. No. 8.7e-19;
Matches 148; Conservative 58; Mismatches 188; Indels 113; Gaps 21;

QY 11 RVLFGWLLGETISGCGYGLWLDDEARTCFRVPWKHFARKDLS-EADARIFKAWAVAR 69
DB 11 RVLFGWLLGETISGCGYGLWLDDEARTCFRVPWKHFARKDLS-EADARIFKAWAVAR 69
14 RVLRLKWLVAQNSCCYGLQWVNGSKLLFYIPWRHATRHGSDQDNTIFKAWAKETGK 73
QY 70 WPPSRGGGPPPEAETAEAGWKTNRFCALRSTRFRVWMLRDSGD-PADPHKVVYALSREL 128
DB 74 YTEG-----VDEADPAKWKANURCALNRSRDLQFYDGRDMPPOPKYKIY----EV 120
QY 129 CWREGP-----GTDQTEAFAAVPPPGPGPPGLAHTHAGLQAPGLPAP 175
DB 121 C-SNGPAPTESQTDYDVLGEEEEEEL-----QRMLPG--LSITEPAL--PGPPNAP 170
QY 176 AGDKGDLQLAQVQSCSLADHLTLTASWGADVPFTKAFCEGQGLPLTFCAGAGGGLPAGEL 235
DB 171 -----YSLPKEDTKWPPALQPPVGL-----GPPVDPDLN 199
QY 236 YGWAETTPSGPQ-----PAALTGEEAAPSPHOAEPLY-----SPSPACTAVQEP 284
DB 200 L-----APPSGNPAGFRQLLPEVLEPGPLASSQPP--TEPLLPDLLLISPHMLPLT----- 247
```

```
QY 285 SPGALDVTIMYKGRVTLQKVVGHP-SCTFLYGPPDPAVRATDPO-----OVAPPS 333
DB 248 ---DLEIKFYQGRAPRTLTISNPGGCRLFYS-----QEAQEQVELFGPVTLEQVRPFS 300
QY 334 PAELPDQKQLRYTBELLRHVAPGLHLELRGQWARRMGKCKVYWEVGPPGSGASPTPA 393
DB 301 PEDIPSKQREYTNQLLDVLRGLILQOGDLYAIRLCQCKVFW--SGPCALAHGSCFN 358
QY 394 CLLPNCCTPIDFRVFQELVVEFRARORRSPRYTYIYLGFGQDLSAGRPKESILVVKL 453
DB 359 -PIQREVTKLFSLEQFLNELILFKQGTNTPTTPEIFFCFGEWPPDVKPREKKLITQV 417
QY 454 EPWLCRVHLEGTQREGVSSLDSDLDL 480
DB 418 VFVAARLLLEMFSGELSSWADSIRLQI 444

RESULT 7
IRF3_MOUSE
ID IRF3_MOUSE STANDARD; PRT; 419 AA.
AC P70671;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interferon regulatory factor 3 (IRF-3).
GN IRF3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Hakem R., Grossman A., Antonio L., Suggs S., Mak T.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1 (BY SIMILARITY).
CC 2- SUBCELLULAR LOCATION: Nuclear.
CC 3- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75839; AAB36924.1; -
CC EMBL; U75840; AAB36925.1; -
CC HSSP; P23906; 2IRF.
CC MGD; MGI:1859179; Irf3.
CC InterPro; IPR001346; IRF.
CC Pfam; PF00605; IRF; 1.
CC PRINTS; PR00267; INTERNEGCT.
CC ProDom; PD002355; IRF; 1.
CC SMART; SM00348; IRF; 1.
CC PROSITE; PS00601; IRF; 1.
CC Transcription regulation; DNA-binding; Nuclear protein; Activator.
KW DNA BIND 7 107 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 419 AA; 46852 MW; 1FF67C4E0FC7F027 CRC64;

Query Match 16.0%; Score 438; DB 1; Length 419;
Best Local Similarity 27.6%; Pred. No. 1.7e-18;
Matches 143; Conservative 56; Mismatches 182; Indels 138; Gaps 17;

QY 6 BRAPRVPLFGEVLLGEISGCGYGLWLDDEARTCFRVPWKHFARKDLS-EADARIFKAWAV 65
DB 2 ETPKPRIL--PWLVSQDLQGLEGVANWLDSETRFRIPWKHGLRQDAQWADFGIFQAWAE 59
QY 66 ARGRWPPSSRGGGPPPEAETAEAGWKTNRFCALRSTRFRVWMLRDSGDPAADPHKVVYALS 125
```

```
Db 60 ASGAYT-----GKDKPDVST-----WKNRFRSALNRKEVRLAADNSKDPYDHPKVEFY 110
QY 126 RELCWREGPGT-----DQTEAEAPAAVPPGPGPGPLAHTAGLQAPGPLPAPAGD 178
Db 111 T-----PGARDPVHUGASPDNGKSSLSHQSLENLPKLP-----DGLILGPLKD 153
QY 179 KGDLLQAOVQOCLADHLLTASWGADPVPTKAPGEGGLPLTACACAGGGLPAGELYGW 238
Db 154 EGSSDLAIV-----SDP----- 165
QY 239 AVETTPSPGPQPAALTTGEAAAPESPHQAEPYLSPPSACTAVQESPGL-----DV 291
Db 166 -SQOLPSP-----NANFNLPAPQ-----ENPLKQLLABEOWEFV 200
QY 292 TIWKGRITVLQKVGHPSCTFLXGPPDPVAVRATDPOQVAFPSP-AELPDQKQLRYTEEL 350
Db 201 TAYRGRQVQQTILFCGGGLVLV--STADMTLPWQVTLDPDEGFLTDKLVEYVGQVL 258
QY 351 RHVAPGLHLELRGPQLWARMGKCKVYWEVG-----GPPGSASPTPACLLPRNCD 401
Db 259 KGLGNGLALWAGQCLWAQRLGHSHAFWALGEELLPDGSRGPDGE-----VHKDKD 309
QY 402 TPIPDFRVFQELVEFRARORRSPRYTIVLFGQDLSAGRPKESLVLVLEPWLVRVH 461
Db 310 GAVFDLRFPVADLIAFMEGSGH-SPRYTLWFCMGEMWPDQDPVWVKRLVMVVKVVTCLKEL 368
QY 462 LEGTOREGVSSDSDLDLCLSSANSI-YDIECFELMEL 499
Db 369 LEMAREGASSLKTVDLHISNQSILTSQYKAYLQDL 407

RESULT 8
IRF6 MOUSE STANDARD; PRT; 467 AA.
ID IRF6 MOUSE AC P97431;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 6 (IRF-6).
GN IRF6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Colon;
RA Grosman A., Mitrucker H.W., Antonio L., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U73029; AAB36714.1; -.
DR HSSP; P23906; 2IRF.
DR MGD; MGI:1859211; Irf6.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTFRNREGFT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 9 111 TRIPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 467 AA; 53106 MW; 68CCAA90680FEDC8 CRC64;
```

```
Query Match 16.0%; Score 436; DB 1; Length 467;
Best Local Similarity 27.7%; Pred. No. 2.5e-18;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 17;

QY 1 MALAPRAAPRVLFGLWLGELISSCYEGQLWIDEARTCFRVPWKHFAR-KDISEADARI 59
Db 1 MALHPR-----RVLKLPWLVAQVDSGLYPLGLIWLHRDSKRFQIPWKHATRHSPQOEENI 56
QY 60 FKAWAVARWRPSSSSGGGPPPEAETAERAGWKTNERCALRSTRFRVLMRDNSG-PADP 118
Db 57 FKAWAVETGKY-----QEGVDDDDPAK-----WKAQLRCLANKSREFNLMYDGTKEVPMNP 107
QY 119 HKVYALSRELWCWREGPGTQTEAEAPAAVPPGPGPGPLAHTAGLQAPGPLPAPAGD 178
Db 108 VKIY-----QVC-----DIPOQG-----SVINPGSTGSAPWD 135
QY 179 KGDLLQAOVQOCLADHLLTASWGADPVPTKAPGEGQGLPL-----TGACAGGGLPAGE 234
Db 136 EKNDVDEDEE-----EDELEQSQHHPVPIQDTFFPLNINGSMPAPASVGVNCSVNCSP--- 188
QY 235 LYGWAVETTPSPGPQPAALTTGEAAAPESPHQAEPYLSPPSACTAVQESPGLDVTIM 294
Db 189 -----ESVMPKTEPLEMEVPPQAPIQFFYSSPELWISSLPMT-----DLDIKFQ 231
QY 295 YKGRITVLQKV-VGHP-SCTFLYGPDP-----AVRATDPOQVAFPSPAELPDQKQLRYT 346
Db 232 YRGEYQVQMTVSNPOGCRIFYGDLGMPDQBELFGVSLQEVKFGPGEHITNEKQLFT 291
QY 347 EELLRHVAPGLHLELRGPQLWARMGKCKVYWEVGPPGSASPTPACLLPRNCDTPIFD 406
Db 292 SKLLDVMDRGLILEVSGHAIYAIRLCQCKVYWSGPCAPSLAEPN-----LIERQKVKLFC 347
QY 407 FRVFFOELVEFRARORRSPRYTIVLFGQDLSAGRPKESLVLVLEPWLVRVHLEGTQ 466
Db 348 LETFLSELIAHKGQIEKQPPFEIYLCFGEEDWPGKLERKLLVQVPIPVVARNIEMFMS 407
QY 467 REGVSSDSDLDLCLSSANSIYDIECFELMEL 499
Db 408 GDFTSRFSDGSVRLQISTPD-INKNIVAQLKQL 439

RESULT 9
IRF6 HUMAN STANDARD; PRT; 467 AA.
ID IRF6 HUMAN AC O14856;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 6 (IRF-6).
GN IRF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grosman A., Mitrucker H.W., Antonio L., Ozato K., Mak T.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


CC the European Bioinformatics Institute. There are no restrictions on its
CG use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF027292; AAB84111.1; --
DR EMBL; AL022398; CAA18545.1; --
DR EMBL; BC014852; AAH14852.1; --
DR HSSP; P23906; 2IRF.
DR Genew; HGNC:6121; IRF6.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 9 111 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 467 AA; 53129 MW; 7E28F5E0F5BA4053 CRC64;

Query Match 15.8%; Score 431; DB 1; Length 467;
Best Local Similarity 27.7%; Pred. No. 4.8e-18;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;

QY 1 MALAPERRAARVLFGEWLLGELISSGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSEADARI 59
DB 1 MALHPR---RVRLKPLWVAQVDSGLYGLIWLHRDSKRFQIPWKHATRHSPQEEENI 56

QY 60 FKAWAVARGWRPSSRGCGPPPEAEAEAGKNTFCALSTRFRVWLRDMSGD-PADP 118
DB 57 FKAWAVETGKY---QEGVDDPPAK-----WKAQLRCALNKSREFNLMYDGTKEVPMNP 107

QY 119 HKVYALSRCLWREGPGTDTAEAPAAVPPPGGPPGPFPLAHTHAGLQAPGLPAPAGD 178
DB 108 VKIY---QVC-----DIPQFG-----SIINFGSTGAPWD 135

QY 179 KGDLLQAVQOQSLADHLHTASGADVPVTKAPGEGQGLPL----TGACAGGPGGLPAGE 234
DB 136 EXD---NDVDEDEDEDE-LDQSHVPIQDTFFPLINGSWAPASVGCNSVGCNSPEAV 191

QY 235 LYGWAVETPSPGQPAALTTGAAAPSPHQAEPYLSPSACTAVQPSFGALDVTIM 294
DB 192 ---W-----PKTEPLEMEVPO--APIQPFYSSPELWISSLPMT-----DLDIRFQ 231

QY 295 YKGRTVLQKV-VGHP-SCTFLYGPPDP-----AVRATDPOQVAPSPAPLDPKOLRYT 346
DB 232 YKGEYGVQMTVSNPOGCRLLFYGDLPMPDQDEELFGVSLQVKFGPEHITNEKQKLF 291

QY 347 EELLRHVAPGLHLELPGPOLWARRMGKCKYVEVGGPPGSGSPSTPACLLPRNCPTPIFD 406
DB 292 SKLLDVMDRGLILEVSGHAIYAIRLCCKVYNSGCPACSLVAPN-----LIERQKKVKLFC 347

QY 407 FRVFFOELVEFRARQRGSPRYTIYLGQODLSAGRPKESLVLVKLEPWLKRVHLEGTO 466
DB 348 LETFLSLDLIAHQKQIEKQKOPPPFIYLCFGEWPDGKPLERKILVQVIVPVARMIYEMS 407

QY 467 REGVSSLDLCLSSANSLVDDIECFMEL 499
DB 408 GDFTRFDSGVSRLQISTPD-IKDNIVAKQL 439

RESULT 10
IRF4_MOUSE
ID IRF4_MOUSE STANDARD; PRT; 450 AA.
AC Q64287; Q60802;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 4 (IRF-4) (Lymphocyte specific interferon
DE regulatory factor) (LSIRF) (NF-EM5) (PU.1 interaction partner)
DE (Transcriptional activator PIP).
GN IRF4 OR SPIP.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C;
RX MEDLINE=95317607; PubMed=7797077;
RA Eisenbeis C.F., Singh H., Storb U.;
RT "Pip, a novel IRF family member, is a lymphoid-specific,
RT PU.1-dependent transcriptional activator.";
RL Genes Dev. 9:1377-1387(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6, and 129/SVJ; TISSUE=Spleen;
RX MEDLINE=95334364; PubMed=7541907;
RA Matsuyama T., Grossman A., Mittrucker H.-W., Siderovski D.P.,
RA Kiefer F., Kawakami T., Richardson C.D., Taniguchi T., Yoshinaga S.K.,
RA Mak T.W.;
RT "Molecular cloning of LSIRF, a lymphoid-specific member of the
RT interferon regulatory factor family that binds the interferon-
RT stimulated response element (ISRE).";
RL Nucleic Acids Res. 23:2127-2136(1995).
CC -|- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
CC WITH PU.1. PROBABLY PLAYS A ROLE IN ISRE-TARGETED SIGNAL
CC TRANSDUCTION MECHANISMS SPECIFIC TO LYMPHOID CELLS.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THE PRESENCE
CC OR ABSENCE OF A GLUTAMIC ACID RESIDUE.
CC -|- TISSUE SPECIFICITY: LYMPHOID CELLS.
CC -|- INDUCTION: NOT INDUCED BY INTERFERONS.
CC -|- SIMILARITY: BELONGS TO THE IRF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U34307; AAA75283.1; --
DR EMBL; U11692; AAA75309.1; --
DR EMBL; U20949; AAA75316.1; --
DR EMBL; U20949; AAA75317.1; --
DR HSSP; P23906; 2IRF.
DR MGD; MGI:1096873; Irfa.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Alternative splicing.
FT DNA_BIND 23 125 TRYPTOPHAN PENTAD REPEAT.
FT VARSELIC 165 165 MISSING (IN ISOFORM 2).
SQ SEQUENCE 450 AA; 51577 MW; 5FD94CA6C453869C CRC64;

Query Match 13.8%; Score 376; DB 1; Length 450;
Best Local Similarity 28.1%; Pred. No. 6.3e-15;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAVAVARGRPSS 74
DB 26 QWLIDQIDSGKYPGLVWNEEKSVFRIPWKHAGKQDYNEEDAAALFKAWLFGKF--- 81

QY 75 RGGGPPPEAEAEAGKNTFCALSTRFRVWLRDMSG-DPADPHKVALSRELWCWREG 133
DB 82 REGIDKDPDPT-----WKTRLRCALNKSNDFBELVRSQIDISDPYKVI----- 127

QY 134 PGTDQTEAEAPAAVPPQGGPPG---FLATHAGLQAPGLPAPAGDKGDLILLQAVQOS 130
Db 128 -----VPEGAKGAKQLTLDQTQWAGHPYPMATP---YGSLLPAQVHNY 169
QY 191 CLADHLLTAGWADPVTAKPAGQEGELPTGACAG-----GGLP-----AGELYGWAVE 241
Db 170 MPPH---DRSW-RDYAPDQSHPEIPYQCPVTFGRHHWQGPSCENGCGVGTGTFYACAPP 226
QY 242 TTPSPG-PQPAALTTGEAAPESPHQAEPLSPSPSACTAVQSPSGALDVTIMYKGRV 300
Db 227 ESQAPGPIPSIRSASALA-----LSDC-----RLHICLYYRDILV 263
QY 301 LQKVVGHP-SCTFLYGPDPDAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLHL 359
Db 264 KELTTTSPGCRISHG---HTYDVSNLDQVLFYP---DNGQRKNIEKLLSHLERGLVL 317
QY 360 ELRGPOLWARMGKCKYVWVGPPGSAS--PSTPACLLPRNCPTPIDFRVPFQELVEF 417
Db 318 WMAPDGLYAKRLCOSRIYD--GPLALCSDRPNK---LERDQTKLFDQQLFSELOVF 371
QY 418 RARQRGSPRYTYLGFQDLSAGRPEKSLVLVLEPWLCR 459
Db 372 -AHGRPAPRFQVTLCEGEEFPDPQ-RQRKLITAHVEPLLAR 411

RESULT 11
IRF4 HUMAN
ID IRF4 HUMAN STANDARD; PRT; 451 AA.
AC Q15306; Q99660;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 4 (IRF-4) (Lymphocyte specific interferon
DE regulatory factor) (LSIRF) (NF-EM5) (Multiple myeloma oncogene 1).
GN IRF4 OR MUM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Spleen;
RX MEDLINE=97079690; PubMed=8921401;
RA Grossman A., Mittrucker H.W., Nicholl J., Suzuki A., Chung S.,
RA Antonio L., Suggs S., Sutherland G.R., Siderovski D.P., Mak T.W.;
RT "Cloning of human lymphocyte-specific interferon regulatory factor
RT (hLSIRF/hIRF4) and mapping of the gene to 6p23-p25.";
RL Genomics 37:229-233(1996).
[2]
RN
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Spleen;
RA Iida S., Butler M., Hatzivassiliou G., Pulivarthi R., Klein B.,
RA Caradine P., Boccadoro M., Nilsson K., Chaganti R.S.K.,
RA Dalla-Favera R.;
RT "Deregulated expression of MUM1 gene by t(6;14) (p25;q32) chromosomal
RT translocation in multiple myeloma.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
CC STIMULATED RESPONSE ELEMENT (ISRE) OF THE MHC CLASS I PROMOTER.
CC BINDS THE IMMUNOGLOBULIN LAMBDA LIGHT CHAIN ENHANCER, TOGETHER
CC WITH PU.1. PROBABLY PLAYS A ROLE IN ISRE-TARGETED SIGNAL
CC TRANSDUCTION MECHANISMS SPECIFIC TO LYMPHOID CELLS.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THE PRESENCE
CC OR ABSENCE OF A GLUTAMIC ACID RESIDUE.
CC -I- TISSUE SPECIFICITY: LYMPHOID CELLS.

CC -I- INDUCTION: NOT INDUCED BY INTERFERONS.
CC -I- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52682; AAC50779.1; -;
DR EMBL; U63738; AAB37258.1; -;
DR EMBL; BC015752; AAH15752.1; -;
DR HSSP; P23906; 2IRF.
DR TRANSFAC; T04929; -;
DR Genew; HGNC:6119; IRF4.
DR MIM; 601900; -;
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
DR PROSITE; PS00601; IRF.1.
DR Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Alternative splicing.
FT DNA_BIND 23 125 TRYPTOPHAN PENTAD REPEAT.
FT VARSPPLIC 165 165 MISSING (IN ISOFORM 2).
FT CONFLICT 300 300 Q -> H (IN REF. 2).
FT CONFLICT 306 306 K -> N (IN REF. 2).
FT CONFLICT 333 333 R -> T (IN REF. 2).
SQ SEQUENCE 451 AA; 51772 MW; 17CD1327C65BF6A CRC64;
Query Match 13.3%; Score 363; DB 1; Length 451;
Best Local Similarity 27.5%; Pred. No. 3.5e-14;
Matches 136; Conservative 73; Mismatches 193; Indels 92; Gaps 24;
QY 16 EWLGLISSGCVGELQWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVGRWPPSS 74
Db 26 QWLIDQIDSGKVPGLVWENEKSIIFRIPWKHAGKQDYNEEDAAALFKANALFKGF---- 81
QY 75 RGGGPPPEAETAERAGKTNFRCAKSTRFRVMLRDNAG-DPADPHKVALSRELWREG 133
Db 82 REGIDKPDPT-----WKRLRCALNKSNDFELVRSOLDISDPKYRIVPEGAK-- 134
QY 134 PGTDQTEAEAPAAVPPQGGPPGFLATHAGLQAPGLP-LPAPAGDKGDLILLQAVQOSCL 192
Db 135 -GAKQLTLEDPMQ-----SMSPYPT-MTTPYPSLPA-----QOVHNYMM 171
QY 193 ADHLLTAGWADPVTAKPAGQEGELPT-----CACAGGPGLPAGELYGWAVET 242
Db 172 PP--LDRSW-RDYVDPQHPPEIPYQCPMTFGRHHWQGPACENGQV-TGTFYACAPPE 227
QY 243 TTPSPG-PQPAALTTGEAAPESPHQAEPLSPSPSACTAVQSPSGALDVTIMYKGRVTL 301
Db 228 SQAPGVPTPEPSIRSASALA-----FSDC-----RLHICLYYRETLVK 264
QY 302 QKVVGHP-SCTFLYGPDPDAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLE 360
Db 265 ELTTSPEGCRISHG---HTYDVSNLDQVLFYP---EDNGQRKNIEKLLSHLERGVVLW 318
QY 361 LRGPOLWARMGKCKYVWVGPPGSAS--PSTPACLLPRNCPTPIDFRVPFQELVEPR 418
Db 319 MAPDGLYAKRLCOSRIYD--GPLALCSDRPNK---LERDQTKLFDQQLFSELOQAF- 371
QY 419 ARQRGSPRYTYLGFQDLSAGRPEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDL 478
Db 372 AHGRSLPRFQVTLCEGEEFPDPQ-RQRKLITAHVEPLLARQLYFYFAQQNSGHFLRGYDL 430
QY 479 DLCLSSANSLYDDI 492
Db 431 PEHISNPEDYHRSI 444

```
RESULT 12
ICSB CHICK STANDARD; PRT; 425 AA.
ID ICSB CHICK STANDARD; PRT; 425 AA.
AC Q90871;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Interferon consensus sequence binding protein (ICSBP).
GN ICSBP1 OR ICSBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95241453; PubMed=7536924;
RA Jungwirth C., Rebbert M., Ozato K., Degen H.J., Schultz U.,
RA David I.B.;
RT "Chicken interferon consensus sequence-binding protein (ICSBP) and
RT interferon regulatory factor (IRF) 1 genes reveal evolutionary
RT conservation in the IRF gene family."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3105-3109(1995).
CC -!- FUNCTION: SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION
CC OF TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON
CC CONSENSUS SEQUENCE (ICS)). PLAYS A REGULATORY ROLE IN CELLS OF THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L39767; AAA62159.1; -
CC HSSP; P23906; 2IRF.
CC InterPro; IPR001346; IRF.
CC Pfam; PF00605; IRF; 1.
CC ProDom; PD002355; IRF; 1.
CC SMART; SM00348; IRF; 1.
CC PROSITE; PS00601; IRF; 1.
CC Transcription regulation; DNA-binding; Activator; Nuclear protein.
CC DNA BIND 9 110 TRYPTOPHAN PENTAD REPEAT.
CC SEQUENCE 425 AA; 49171 MW; 92BB8A9B77024EB2 CRC64;

Query Match 12.4%; Score 338; DB 1; Length 425;
Best Local Similarity 24.3%; Pred. No. 8.7e-13;
Matches 121; Conservative 77; Mismatches 168; Indels 132; Gaps 19;

QY 16 EWLGLBISGCGYGLQWLDSEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPSS 74
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 12 QWLIEQIDSQYPLGLWENEEKTMFRIPWKHACKQDYNQEVDSIFKAWAVRGKF- 67
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY 75 RGGGPPPEAETARAGKWNFRCALSTRFRVFLMRNSG-DPADPHKVVAL--SRELCHWR 131
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 68 -----KEGDKABPATWKTALCALNKSPPDFEETDRSOLDISEPKYVRIVPEEQKCK 121
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY 132 EGPQ-----TDQTEAE-APAAVPPQGGPPGPFPLAHTAGLQAPGLPAPAGDKGDLILQ 185
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 122 IGVNGSSULTVGDMDGSPDAIDLMKEPP----- 151
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY 186 AVQOCSLADHLTASMGADVPVTKAPGEGQGLPLTCACAGGPGPLPAGELYGWAETPS 245
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 152 -----CVDEYL-----GIIKESPPQES-----TCRNP-----IPDWMQ-QPS 185
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY 246 PGQPAALTGEEAAAPESHQABEPLYSPPSACTAVQEPSG--ALDVTIMYKGRVTLQK 303
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
```



```

RESULT 15
ICSB_HUMAN STANDARD; PRT; 426 AA.
AC Q02556;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon consensus sequence binding protein (ICSBP).
GN ICSBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung, and Monocytes;
RX MEDLINE=93094284; PubMed=1460054;
RA Weisz A., Marx P., Sharf R., Appella E., Driggers P.H., Ozato K.,
RA Levi B.-Z.;
RT "Human interferon consensus sequence binding protein is a negative
RT regulator of enhancer elements common to interferon-inducible
RT genes.";
RL J. Biol. Chem. 267:25589-25596(1992).
RN [2]
RP REVISIONS.
RA Schmidt M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION
CC OF TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON
CC CONSENSUS SEQUENCE (ICS)). PLAYS A NEGATIVE REGULATORY ROLE IN
CC CELLS OF THE IMMUNE SYSTEM.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN LYMPHOID TISSUES.
CC -!- INDUCTION: BY INTERFERON GAMMA.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.

```

Search completed: June 18, 2003, 12:47:06
Job time : 17 secs

```

134 PGTDTAEAPAAVPPQGGPGLPAPAGKGDLLQAVQSC-- 191
113 --VPBEEQCKLGV-----ATAG-----CVNEVTEMECGR 140
192 -LADHLLTASWGADPVPTKAPGEGGLPLTGACAGGPGLPAGELYGWAIVETTPSPGPQP 250
141 SEIDELIKE-----PSVDDYMGWIKRSPSP----- 165
251 AALTGEEAAAPESPHQAEPLYSPSACTA-----VOEPSPGA----- 288
166 -----PEACRSQLLPDWAQPGSTGVPLVTGYTTYDAHSA 201
289 ---LDVTIMYKRTVLQKVGHPS-SC-----TLYGPPDPDAVRATDQQAFFPS 333
202 PSQWISFYGGKLVGQATTTCTEGCRSLSPGLPGTKLYGP-----EGLELYRFP 254
334 PAELPDOKOLRYTEELLRHVAPGLHLELRGPOLWARRMGCKVYWEVGGPPGASASTPA 393
255 ADAIPSERQQRQVTRKLFHLERGVLLHSSRQGVFVKRLCQGRVFCSCNAVVKGRNK-- 312
394 CLLPRNCDTPIFDPRVFFQELVEFRARQRRGSPRYTYILGPGQDLSAGRPKEKSLVLVK 453
313 --LERDEVVQVFDTSQFFRELQGFYNSQGR-LPDGRVVLFCGEEPPDMLRSLKILVQI 369
454 EPWLCRVHLEGTQREGVSSLDSSDL-----DLCLSSANSLY 489
370 E----QLYVRQLAEAGKSCGAGSVMAPEEPPDQVFRMFPDICASHORSFF 418

```

```

Query Match 11.3%; Score 308.5; DB 1; Length 426;
Best Local Similarity 21.8%; Pred. No. 4.2e-11;
Matches 116; Conservative 71; Mismatches 161; Indels 185; Gaps 18;
QY 16 EMLLGRISGCGYGLQWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74
DB 12 QWLIEQIDSSMYPLGIWENEKSMFRIPWKCHAGQDYNQEVDAIFKAWAVFKGK- 67
QY 75 RGGGPPPEAETAEACWKNFRCALRSTRFVWMLRNSG-DPADPHKVVALSRELWREG 133
DB 68 -----KEGDKAEPATWTKRLCALNKSPDFEEVTRSDISDPYKYRI----- 112

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:43:22 ; Search time 36 Seconds
(without alignments)
2878.936 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAAPRVLFGEWLLG.....SANSLYDDIECFLEMELEQPA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1116	40.9	210	4	Q9UE79
2	778	28.5	491	13	Q90ZD4
3	603	22.1	128	4	Q9HB64
4	440	16.1	459	13	O57578
5	430	15.7	467	6	Q9N136
6	428	15.7	467	11	Q91VD0
7	420	15.4	467	6	Q8WQ4
8	364.5	13.3	452	4	Q96GL3
9	364	13.3	445	13	Q98TX7
10	362	13.3	409	13	Q98TX6
11	344	12.6	440	4	Q99419
12	321	11.8	416	13	Q90W10
13	281.5	10.3	330	11	Q90ZL7
14	194	7.1	215	11	Q924T6
15	183	6.7	1334	16	Q9RKR9
16	177.5	6.5	296	13	Q98TA1

ALIGNMENTS

RESULT 1

Q9UE79 PRELIMINARY; PRT; 210 AA.
ID Q9UE79;
AC Q9UE79;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2001 (Tremblrel. 16, Last annotation update)
DE Putative interferon regulatory factor 7C.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND BREAST;
RX MEDLINE=97459673; PubMed=9315633;
RA Zhang L., Pagano J.S.;
RT "IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency";
RT Mol. Cell. Biol. 17:5748-5757(1997).
DR EMBL; U53832; AAB80691.1; -;
SQ SEQUENCE 210 AA; 23592 MW; 45895671CEAA18F5 CRC64;

Query Match 40.9%; Score 1116; DB 4; Length 210;
Best Local Similarity 99.0%; Pred. No. 9.5e-70;
Matches 208; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 294 MYKGRVTLQKVGVGHPSCFGLYGPDPVAVRATDPQVAFPSPAELPDQKQLRYTEELLRHV 353
DB 1 MYKGRVTLQKVGVGHPSCFGLYGPDPVAVRATDPQVAFPSPAELPDQKQLRYTEELLRHV 60
QY 354 APGLHLELRGPQLWARRMGKCKVYVEVGGPPGSPSPSTACLLPRNCDTPIFDFRVFFQE 413
DB 61 APGLHLELRGPQLWARRMGKCKVYVEVGGPPGSPSPSTACLLPRNCDTPIFDFRVFFQE 120
QY 414 LIVEFRARQGRGSPRTIYILGFCODLSAGRPKESLVLVKLEPWLCRVHLEGTOREGVSSL 473
DB 121 LIVEFRARQGRGSPRTIYILGFCODLSAGRPKESLVLVKLEPWLCRVHLEGTOREGVSSL 180

```
QY 474 DSDLDLCLSSANSYDDIECFLEMELEQPA 503
DB 181 DSSLSLCLSSANSYDDIECFLEMELEQPA 210

RESULT 2
Q90ZD4 PRELIMINARY; PRT; 491 AA.
AC Q90ZD4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Interferon regulatory factor-3.
GN IRF-3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA May D.L., Grant C.E., Deeley R.G.;
RT "Cloning and Promoter Analysis of the Chicken Interferon Regulatory
RT Factor-3 Gene.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268079; AAK58583.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.
DR ProDom; PD002355; IRF.
DR ProSITE; PS00601; IRF.
DR SMART; SM00348; IRF.
DR NON_TER 128 AA; 14363 MW; 90E0C80A2624B1B4 CRC64;
SQ SEQUENCE 491 AA; 54440 MW; 90E0C80A2624B1B4 CRC64;

Query Match 28.5%; Score 778; DB 13; Length 491;
Best Local Similarity 38.5%; Pred. No. 5.5e-46;
Matches 208; Conservative 59; Mismatches 181; Indels 92; Gaps 17;

QY 2 ALAPRAAPRVFGEWLLGHSICGVEGLQWLDEARTCFRVPWKHFARKDLSADARIFKAWA 61
DB 3 ALDSEGAQKLRFGPMLNVASSLGRLCWIDPDRRIPIPKHNARKDVTSSDVEIFK 62
QY 62 AWAVAGRWPPSSRGSGGPPPEATAERAGWKTFRFCALRSTRFRVWMLRDNSGDPADPHKV 121
DB 63 AWAKASGRY-----EGNAEDPAKWKTFRFCALRSTRFMFLLEDKSKNDPDPHKV 111
QY 122 YALSRELWCWEGPGTQDTQAEAPAAVPPPO.-GGP-----PGPPLAHTHAGLO--- 167
DB 112 YAVA-----SGVPNDRGSGPVAGALQQQPQLLNHHDLAENTPT 152
QY 168 -AFGPLPAPAGDKGL-LLQAVQSCSLADHLLTASWGADVPVTKAPGEG--QEGPLPTG- 222
DB 153 DSTEGVAAALATQVLDLQSLVQHNCISAL-----GSQPTLWAHTGDALPEDALLPGQ 207
QY 223 -ACAGGGLPAGLYGH-AVETTPSGPOP-----AALTGEAAAPSPHQAEP 269
DB 208 DGLPLGP-----QFQDMRQLKEEPLLLGNQPLTGGCGQDGAGALPVSEECALPAPSPAE 262
QY 270 YL-----SPSPSACTAVQEPSPGALDVTIMYKGRVQLQKVGHPSCTFLYGPDPAPVRATD 325
DB 263 LLFQSANPAPPAGDIGGLPPLDITYYRGKVVQEYDSDRCVLAYQPLDPAV--AE 320
QY 326 PQQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLELPGQLMARRMGKCKVYWEV----- 380
DB 321 QRLVLFPSPASLPDRPQRRYTNILLE--VAGLRLEQRAGQLLATRLKCKVFWALSQOLE 378
QY 381 GGPFGSASPSPACILPRNCDTIFDFRVFFQELVEFRAPQRGRSPRYTIYLFQGDLSA 440
DB 379 GGEP-----PLNLLHRDQETTFIDFRVFCTELDRFDRSRRSPDFTIFLCFGQCFSS 431
QY 441 GRPKESLVLVKLEPMLCRVHLEGTQREGVSSDSDLDLCLSSANSYDDIECFLEMELE 500
DB 432 TKPKESKLLVLKLVQPCFYWYQVQVQGGASSLNSGNVLSQLSDSNFLFLIEQYHMQTD 491
```

```
RESULT 3
Q9HB64 PRELIMINARY; PRT; 128 AA.
AC Q9HB64;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Interferon regulatory factor-7H (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20493557; PubMed=10924517;
RA Lu R., Au W.-C., Yeow W.-S., Hageman N., Pitha P.M.;
RT "Regulation of the promoter activity of interferon regulatory factor-7
RT gene. ACTIVATION BY INTERFERON AND SILENCING BY HYPERMETHYLATION.";
RL J. Biol. Chem. 275:31805-31812(2000).
DR EMBL; AF277159; AAG30003.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.
DR PRINTS; PR00267; INTERNRGEGFCT.
DR ProDom; PD002355; IRF.
DR SMART; SM00348; IRF.
DR ProSITE; PS00601; IRF.
DR NON_TER 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;
SQ SEQUENCE 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;

Query Match 22.1%; Score 603; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.5e-34;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVLFGEWLLGHSICGVEGLQWLDEARTCFRVPWKHFARKDLSADARIFKAWA 66
DB 20 RAAPRVLFGEWLLGHSICGVEGLQWLDEARTCFRVPWKHFARKDLSADARIFKAWA 79
QY 67 RGRWPPSSRGSGGPPPEATAERAGWKTFRFCALRSTRFRVWMLRDNSGDP 115
DB 80 RGRWPPSSRGSGGPPPEATAERAGWKTFRFCALRSTRFRVWMLRDNSGDP 128

RESULT 4
Q57578 PRELIMINARY; PRT; 459 AA.
AC Q57578;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE XIRF-6 protein.
GN XIRF-6.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086218; PubMed=9426249;
RA Hatada S., Kinoshita M., Takahashi S., Nishihara R., Sakumoto H.,
RA Fukui A., Noda M., Asashima M.;
RT "An interferon regulatory factor-related gene (xirf-6) is expressed in
RT the posterior mesoderm during the early development of Xenopus
RT laevis.";
RL Gene 203:183-188(1997).
DR EMBL; D86492; BAA24349.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.
DR PRINTS; PR00267; INTERNRGEGFCT.
DR ProDom; PD002355; IRF.
DR SMART; SM00348; IRF.
```



```
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 459 AA; 52475 MW; 306DA77C6586871C CRC64;

Query Match 16.1%; Score 440; DB 13; Length 459;
Best Local Similarity 27.1%; Pred. No. 1.1e-22;
Matches 140; Conservative 74; Mismatches 197; Indels 106; Gaps 19;

QY 1 MALAPERAAAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSADARI 59
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 1 MAHPR-----RVRLKPWLVAQVDSGLYPLGLIWLHRDSKRFQIPWKHATRHSPQEEENTI 56
QY 60 FKAWAVAGRWPPSSRGSGGPPPEAEATBAEAGKTNFRCALRSTRFVMLRDNNGD-PADP 118
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 57 FKAWAVETGKYOG-----ADEPDAKWAQALRCALNKSREFKLMYDGTKEVPMNP 107
QY 119 HKVYALSRELWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGD 178
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 108 VKIY-----EVC-----DIPQSG-----SIINPGSGVPWD 135
QY 179 KGDLLQAVQOCLADHLLTASWGADPVPTKAP-----GEGQGLPLTGACAGGGLPA 232
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 136 DDDFEADLNQ-----QNHVPISEPFNCLINDSPIGSSSTGCT-----176
QY 233 GELYGWAIVETTPSGQPAAL--TTGEAAAPSPHQAEPLSPSACTAVQEPSPGALD 290
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 177 -----PEQTWPKTEPQEMVEPPTSGPADFFSP---EMWISLEMT-----DLE 217
QY 291 VTIMYKGRVLOKV-VGHP-SCTFLYGGPPDPVAVRATD-----PQVAPPSPPAELPDQKQ 342
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 218 IQFYRGKEMGQTMVSNPQGCRLFYDGLGPNQEELEFGPITLQVREPGTEQIVNEKQ 277
QY 343 LRYTELLRHVAPGLHLELRGPOLWARRMGKCKVYVEGPPGSGASPSPTPACLLPRNCDT 402
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 278 KLFTSKLLDVMRGILLEVSHAIYAIRLCQCKVYWG---SGCSPSPITPN-FIERQKRV 333
QY 403 PFDPRVFQELVEFRARQRSPRYTIYLGFGQDLSAGRPKEKSLVVKLEPWLCRVHL 462
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 334 KLFCVETFLSLDLSHQKGIITKQPPPEIYLCFGEWPDGKQERKLIIVQIIPVARMII 393
QY 463 EGTQREGVSSLDSDLDCLSSANSYDDIECFMEL 499
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 394 EMFTGDSTRSFGSGIRLQISIPD-IKDNIVSHLKH 429

RESULT 5
Q9N136 PRELIMINARY; PRT; 467 AA.
ID Q9N136
AC Q9N136
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Interferon regulatory factor 6.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=15 DAY PREGNANT UTERUS;
RA Choi Y., Spencer T.E., Bazer F.W.;
RT "Cloning and Analysis of Ovine IRF-6.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228446; AAF34782.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERFERGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 467 AA; 52970 MW; 21E04F749844D88F CRC64;

Query Match 15.7%; Score 430; DB 6; Length 467;
Best Local Similarity 27.3%; Pred. No. 5.7e-22;
Matches 141; Conservative 72; Mismatches 209; Indels 94; Gaps 20;

QY 1 MALAPERAAAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSADARI 59
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 1 MAHPR-----RVRLKPWLVAQVDSGLYPLGLIWLHRDSKRFQIPWKHATRHSPQEEENTI 56
QY 60 FKAWAVAGRWPPSSRGSGGPPPEAEATBAEAGKTNFRCALRSTRFVMLRDNNGD-PADP 118
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 57 FKAWAVETGKY-----QEGVDDPDPAR-----WKAQLRCALNKSREFNLMYDGTKEVPMNP 107
QY 119 HKVYALSRELWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLP-----173
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 108 VKIY-----QVC-----DIPQGGSIINP-----GSTGSAFWDEKD 138
QY 174 --APAGDGDLLQAVQOCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGGLP 231
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 139 NDVDEDEDELDQSHVPIQDTFFLNINGSPI---APGS-----VGNCSVGNCS 188
QY 232 AGELYGWAIVETTPSGQPAALTTGEAAAPSPHQAEPLSPSACTAVQEPSPGALDV 291
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 189 EAV-----W-----PKTEPLEMEVPQ--APIQPFYSSPELWISSLEMT-----DLDI 228
QY 292 TIMYKGRVLOKV-VGHP-SCTFLYGGPPD-----AVRATDPQVAPPSPPAELPDQKQ 343
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 299 KFQYRGKEMGQTMVSNPQGCRLFYDGLGPNQEELEFGPVSLEQVKGFPGEHITNEKOK 288
QY 344 RYTELLRHVAPGLHLELRGPOLWARRMGKCKVYVEGPPGSGASPSPTPACLLPRNCDT 403
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 289 LFTSKLLDVMRGILLEVSHAIYAIRLCQCKVYWGSCAPSLVAPN---LIERQKKVK 344
QY 404 IPDFRVFQELVEFRARQRSPRYTIYLGFGQDLSAGRPKEKSLVVKLEPWLCRVHL 463
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 345 LFCLETFSLDLSIAHKQGIETKQPPPEIYLCFGEWPDGKQERKLIIVQVIVVAGMIYE 404
QY 464 GTQREGVSSLDSDLDCLSSANSYDDIECFMEL 499
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 405 MFGDFTSRFSGSVRLQISTPD-IKDNIVAQLKQ 439

RESULT 6
Q91VD0 PRELIMINARY; PRT; 467 AA.
ID Q91VD0
AC Q91VD0
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE BM282D4.4 (interferon regulatory factor 6).
GN IRF6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tromans A.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL365322; CAC42184.1; -.
DR EMBL; BC008515; AAH08515.1; -.
DR MGD; MGI:1859211; Irfg.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 467 AA; 53110 MW; 1C564BC8D79C5259 CRC64;

Query Match 15.7%; Score 428; DB 11; Length 467;
Best Local Similarity 27.5%; Pred. No. 7.8e-22;
```

```
Matches 141; Conservative 72; Mismatches 212; Indels 88; Gaps 17;
Qy 1 MALAPERAAARVLFGWLLGEISSGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSEADARI 59
Db 1 MALHPR-----RVRLKWLVAQVDSGLYGLIWLHRDSKRFOIPWKHATRHSPOOEENTIFKAWAETGK 56
Qy 60 FKAWAVARGSPSSRGSGGPPPEAETAAERAGWKTNRFCALRRTRRRFVMLRDNSSGD-PADP 118
Db 57 FKAWAVETGKY---QEGVDDPDPAK-----WKAQLRCLANKSRFNLMDYDGTKEVPMNP 107
Qy 119 HKVYALSRELWCWREGPGTQTEAEAPAAVPPPOGGPPGFLAHTHAGLQAPGPPPLAPAGD 178
Db 108 VKIY-----QVC-----DIPQTQ-----SVINPGSTGSAPWD 135
Qy 179 KGLLLQAVQOOSCLADHLLTASGADPVPTKAPGEGEGLPL-----TGACAGGPGLPAGE 234
Db 136 EKONDVEDEE-----EDELQSOHHVPIQDTFFPLNINGSPIAPASVGNCSVGNCSGP--- 188
Qy 235 LYGWAVETTPSPGPQPAALTTGAAAPESPQAEPLSPSPSACTAVQEPSGALDVTIM 294
Db 189 -----ESVMPKTEPLEMEVPPQAPIQPFYSSPELWISSLPMT-----DLDIKFQ 231
Qy 295 YKGRVTLQKV-VGHP-SCTFLYGPDP-----AVRATDPOQVAFSPAPBLDPKQOLRYT 346
Db 232 YRGKEYGQITVTSNPOGCRLLFYGLGMPDQEBELFGPVSLEQVKFGPEHITNEKQKLT 291
Qy 347 EELLRHVAPGLHLELGPQLWARMGKCKYVWEGGPPGSPASSTPACLLPRNCDTPIED 406
Db 292 SKLVDVMDRGLLEVSGHAIYAIRLCOCKYVWEGGPPGSPASSTPACLLPRNCDTPIED 347
Qy 407 FRVFFOBLVFRARQRGSPRYTIYLFQGDLSAGRPKESLVLVKLEPWLCLRVHLEGTQ 466
Db 348 LETFLSELIAHQKQIEKQPPFIYLCFGEWPDGKLERKLIIVQVIVVARMIVEMFS 407
Qy 467 REGVSSLDSDLDLCLSSANSLYDDIECFMEL 499
Db 408 GDFTRSPDGSVRLQISTPD-IKDNIVAQLKQL 439

RESULT 7
Q8WNQ4 PRELIMINARY; PRT; 467 AA.
AC Q8WNQ4;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Interferon regulatory factor 6.
GN IRF6.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ILEUM;
RA Farber C.R., Raney N.E., Ernst C.W.;
RT "Characterization of the porcine IRF6 gene: cDNA cloning, expression
RT analysis and chromosomal localization."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327368; AAL37429.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 467 AA; 53047 MW; 4AB757DA8013A3C2 CRC64;

Query Match 15.48; Score 420; DB 6; Length 467;
Best Local Similarity 27.48; Pred. No. 2.8e-21;
Matches 138; Conservative 71; Mismatches 210; Indels 84; Gaps 19;
Qy 11 RVLFGEWLLGEISSGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSEADARIFKAWAVARGR 69
```

```
Db 7 RVRLKWLVAQVDSGLYGLIWLHRDSKRFOIPWKHATRHSPOOEENTIFKAWAETGK 66
Qy 70 WPPSSRGSGGPPPEAETAAERAGWKTNRFCALRRTRRRFVMLRDNSSGD-PADPHKVYALSREL 128
Db 67 Y---QEGVDDPDPAK-----WKAQLRCLANKSRFNLMDYDGTKEVPMNPVKIY---QV 113
Qy 129 CWREGPGTQTEAEAPAAVPPPOGGPPGFLAHTHAGLQAPGPPPLAPAGDKGDLLLQAVQ 188
Db 114 C-----DIPQOG-----SVINPGSTGSAPWDEK---NDVD 142
Qy 189 QSCLDHLLTASGADPVPTKAPGEGEGLPLTGA---CAGGPGLPAGELYGWAVETTP 244
Db 143 DEDEEDE-LDQSOHHVPIQDTFFPLNINGSPIAPASVGNCSVGNCSPEAV---W----- 192
Qy 245 SPGPQPAALTTGAAAPESPQAEPLSPSPSACTAVQEPSGALDVTIMYKGRVTLQKV 304
Db 193 -KPAEPLMEVPO--APIQPFYSSPELWISSLPMT-----DLDIKFQYRKEYGQTM 241
Qy 305 -VGHP-SCTFLYGPDP-----AVRATDPOQVAFSPAPBLDPKQOLRYTEELLRHVAPG 356
Db 242 TVSNPOGCRLLFYGLGMPDQEBELFGPVSLEQVKFGPEHITNEKQKLTFLDVMDRG 301
Qy 357 LHLELGPQLWARMGKCKYVWEGGPPGSPASSTPACLLPRNCDTPIEDFRVFFOBLVE 416
Db 302 LILEVSGHAIYAIRLCOCKYVWEGGPPGSPASSTPACLLPRNCDTPIEDFRVFFOBLVE 357
Qy 417 FRARQRGSPRYTIYLFQGDLSAGRPKESLVLVKLEPWLCLRVHLEGTQREGVSSLDSS 476
Db 358 HOKGQIEROPFFFIYLCFGEWPDGKLERKLIIVQVIVVARMIVEMFSGDFTRSPDGS 417
Qy 477 DLIDLCLSSANSLYDDIECFMEL 499
Db 418 SVRLQISTPD-IKDNIVAQLKQL 439

RESULT 8
Q96GL3 PRELIMINARY; PRT; 452 AA.
AC Q96GL3;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Similar to interferon regulatory factor 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009395; AAH09395.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 452 AA; 49120 MW; 96B059A028AE6B64 CRC64;

Query Match 13.38; Score 364.5; DB 4; Length 452;
Best Local Similarity 27.48; Pred. No. 1.8e-17;
Matches 113; Conservative 43; Mismatches 153; Indels 103; Gaps 16;
Qy 10 PRVLFGWLLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARGR 69
Db 6 PRIL--PWLVSQDLQGLEGVAVWVWVNSRTRFRFPWKHGLRQDAQQEDFGIFQAWAEATGA 63
Qy 70 WPPSSRGSGGPPPEAETAAERAGWKTNRFCALRRTRRRFVMLRDNSSGDPAHPKVYALSREL 129
Db 64 YVP-----GRDKDLPDTP-----WKRNFSAFNKRGELRLAEDRSKDPHDPKHIFEV---- 110
Qy 130 WREGPGTQTEAEAPAAVPPPOGGPPGFLAHTHAGLQAPGPPPLAPAGDKGDLLLQAVQ 189
```

```
Db 111 -NSGVG-----DFSQDPTSDPTNGG-----GSTD-----TQ 136
QY 190 SCLADHLITASGADVPPTKAPGEGGGLPLTACACAGGGLPAGELYGWAVETTPSPGPQ 249
Db 137 EDILDELL-GNNVLAPL-----DFGP-----SLAVAPPCPQ 169
QY 250 PAALTGTEAAAPESPHQAEPPYLSPPSACTAVOEP-SPGALDVTIMYKGRTVLQKVGHP 308
Db 170 PLRSPSLDNPT-----PPNLGSPENPLKRLLVPGEEWFEFVAFYRGQVFOQTISCP 223
QY 309 SCTFLYGGPDPAVRATDPOQVAFPSA-ELPDQKQLRYTEELRLHVAPGLHLELGPOLW 367
Db 224 BGLRIVG-SEVGRITLPGWVTLTPDGMSLTDGKWSYVRHVLSCLGGGLAWRAGQWLW 282
QY 368 ARRMGCKKYWEVG-----GPPGASPTPACLLPRNCDTPIFDPRVF 410
Db 283 AQRLGHCHTYWAVSEBLLPNSGHGPDG-----VPKDEGGVFDLGP 325

RESULT 9
Q98TX7
ID Q98TX7 PRELIMINARY; PRT; 445 AA.
AC Q98TX7
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Interferon regulatory factor 4.
GN IRF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RX MEDLINE=21424622; PubMed=11533227;
RA Hrdlickova R., Nehyba J., Bose H.R. Jr.;
RT "Interferon Regulatory Factor 4 Contributes to Transformation of v-
Rel-Expressing Fibroblasts.";
RL Mol. Cell. Biol. 21:6369-6386(2001).
DR EMBL; AF203331; AAK08198.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 445 AA; 51120 MW; 7CAE7BDF96780432 CRC64;

Query Match 13.3%; Score 364; DB 13; Length 445;
Best Local Similarity 26.4%; Pred. No. 1.9e-17;
Matches 130; Conservative 71; Mismatches 201; Indels 90; Gaps 20;

QY 16 EWLLEISGCGEGQLWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGWRPPSS 74
Db 22 QWLIDQIDSGKYPGLWENDEKSI FRIPWKHAGKQDYNEEDAALFKAWALFKGKF---- 77
QY 75 RGGGPPPEAETAEAGWKTNFRCALRSTRFRVWLKDNNG-DPADPHKYVALSRELCLWREG 133
Db 78 REGIDKPDPT-----WKTRLCALNKSNDFFELVERSOLDISDPYKYVRIVPEGA-KKG 131
QY 134 PGTDQTEAEPAAVPPQGGPGPLAHTHAGLQAPGLP-----APAGDKGDLILL 184
Db 132 AKONGMEEQPLMNHPPFITSP-----YTSLPSPQVNYVPHERNWREFAPEQHPDIPY 185
QY 185 QAVQSCSLADHLITASGADVPPTKAPGEGGGLPLTACACAGGGLPAGELYGWAVETTP 244
Db 186 QCAS-----VFFAARGHHWQPGCENGCO-----VTGTFVACAPPESQ 223
QY 245 SPG-PQPAALTTGEAAAPESPHQAEPPYLSPPSACTAVQEPSPGALDVTIMYKGRTVLQK 303
```

```
Db 224 TPGIPDIESIRSGEALA-----LSDC-----RLHICLYTREMVLKVEV 260
QY 304 VVGHF-SCTFLYGGPDPAVRATDPOQVAFPSAELPDQKQLRYTEELRLHVAPGLHLELR 362.
Db 261 TTSSPEGCRISQ---QSVEVSSLEQVIFPYE---EDNSQRKNIEKLLSHLGRGVILWMA 314
QY 363 GPQLWARMGCKKYWEVGPPGPGSAS--PSTPACLLPRNCDTPIFDPRVFFOELVEFRAR 420
Db 315 PDGLYAKRLCOSRIYWD--GPLALCSDRPNK-----LERDQTKLFDTOQLAELOQAF-AH 367
QY 421 QRRGSPRYTIVLGFQDLSAGRPKEKSLVVKLEPWLRCRVHLEGTQREGVSLDSSDLQ 480
Db 368 HCRPLRPYQVALLCFGEFPDPQ-RORKLITAHVEPMFAQLYFFAQNSCHLLRGVLDLPE 426
QY 481 CLSSANSLYDDI 492
Db 427 LMTSPEDYHRSI 438

RESULT 10
Q98TX6
ID Q98TX6 PRELIMINARY; PRT; 409 AA.
AC Q98TX6
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Interferon regulatory factor 4 delta86.
GN IRF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RX MEDLINE=21424622; PubMed=11533227;
RA Hrdlickova R., Nehyba J., Bose H.R. Jr.;
RT "Interferon Regulatory Factor 4 Contributes to Transformation of v-
Rel-Expressing Fibroblasts.";
RL Mol. Cell. Biol. 21:6369-6386(2001).
DR EMBL; AF203332; AAK08199.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 409 AA; 47489 MW; 0187B3782B62480E CRC64;

Query Match 13.3%; Score 362; DB 13; Length 409;
Best Local Similarity 25.0%; Pred. No. 2.4e-17;
Matches 123; Conservative 65; Mismatches 178; Indels 126; Gaps 18;

QY 16 EWLLEISGCGEGQLWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGWRPPSS 74
Db 22 QWLIDQIDSGKYPGLWENDEKSI FRIPWKHAGKQDYNEEDAALFKAWALFKGKF---- 77
QY 75 RGGGPPPEAETAEAGWKTNFRCALRSTRFRVWLKDNNG-DPADPHKYVALSRELCLWREG 133
Db 78 REGIDKPDPT-----WKTRLCALNKSNDFFELVERSOLDISDPYKYVRI----- 123
QY 134 PGTDQTEAEPAAVPPQGGPGPLAHTHAGLQAPGLPAGDKGDLILLQAVQSCSLA 193
Db 124 -----VPEGAKKGAKQMSMEQPLM 143
QY 194 DHLITASGADVPPTKAPGEGGGLPLTACACAGGGLPAGELYGWAVETTPSPGPQPAAL 253
Db 144 NH-----PPFITSP-----YTSLPSPQVNYM 164
QY 254 TTGE-----AAAPESPHQAEPPYLSPS-PSACTAVQEPSPGA-----LDVTIMYKGRTVLQK 303
```

```
Db 165 VPERNWRFAPOPHDIPYQCASFVFAARGHHWGFCGENDCRHLHICLYYREMLVKEV 224
Qy 304 VVGHPP-SCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLELR 362
Db 225 TTSSPGCRISQ---QSVEVSSLEQVIFYP---EDNSQKNIEKLLSHLGRGVILWMA 278
Qy 363 GPQLMARMGKCKVYEVGGPGSAS--PSTPACLLPRNCDTPIFDPRVFQBELVFRAR 420
Db 279 PDGLYAKRLCQSRIYWD--GFLALCSDRPNK---LERDQCKLFDQQLAELQAF-AH 331
Qy 421 QREGSPRYTYILGFGQDLSAGRPKEKSLVLKLEPWLRCVHLEGTQBEGVSSLDSDLDL 480
Db 332 HGRPLPRYQVLCFGEFPDPQ-RQRKLITAHVEMPFAQLYFYAQONGSHLLRGYDLPE 390
Qy 481 CLSSANSYDDI 492
Db 391 LMTSPDYHRSI 402

RESULT 11
Q99419
ID Q99419 PRELIMINARY; PRT; 440 AA.
AC Q99419
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ICSAT transcription factor (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96239482; PubMed=8657101;
RA Yamagata T., Nishida J., Tanaka T., Sakai R., Mitani K., Yoshida M.,
RA Taniguchi T., Yazaki Y., Hirai H.;
RT "A novel interferon regulatory factor family transcription factor,
RT ICSAT/Pip/LSIRF, that negatively regulates the activity of interferon-
RT regulated genes.";
RL Mol. Cell. Biol. 16:1283-1294(1996).
DR EMBL; D78261; BAAL1335.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERFERGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
FT NON_TER
SQ SEQUENCE 440 AA; 49804 MW; 6A04159FAFBA2701 CRC64;

Query Match 12.6%; Score 344; DB 4; Length 440;
Best Local Similarity 28.1%; Pred. No. 4.6e-16;
Matches 124; Conservative 60; Mismatches 161; Indels 96; Gaps 21;

Qy 16 EWLIGISSCYEGLOWLDEARTCFRPWKHFARKDL-SEADARIFKAWAVARGRPPSS 74
Db 57 QWLIDQIDSKYPLGWENEKESIFRIPWKHAGCKQDYNEEDAAFKAWLFGKF----- 112
Qy 75 RGGGPPPEATBAGKWNFRCALSTRFRFVMLRNSG-DPADPHKVVALSRELWCWG 133
Db 113 REGIDKPDPT-----WKTLRCLANKNSDFEELVERSQIDISDPYKVRIVPEGAK-- 165
Qy 134 PGTDQTEAEAP--AAVPPPGGPGPL-AHTHAGLQAPGLPAPAGDKGDLILLQAVQOS 190
Db 166 -GAKQLTLEDPQMSGHPTMTTPYSLPAQVNYMPP----- 203
Qy 191 CLADHLITASWGADPVPTKAPGEGQELPLT-----GACAGGPGLPAGELYGWAV 240
Db 204 -----LDRSW-RDYVPDQPHPEIPYOCPTWFGRGHHWQGPACENGCV-TGTFYACAP 255
Qy 241 ETTFSFG-POPAALTGTGAAAAPSPHQAEPYLSFSPSACTAVQBPFGALDVTIMYKGR 299
```

```
Db 256 PESQAPGVPTSPSIRSAEALA-----FSDC-----RLHICLYIREIL 292
Qy 300 VLQKVVGHP-SCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLH 358
Db 293 VKELTSSPEGCRISG---HTYDASNLQVLFYP---EDNQKNIEKLLSHLGRGVV 346
Qy 359 LEIRGQQLWARRMGKCKVYEVGGPGSAS--PSTPACLLPRNCDTPIFDPRVFQBELVE 416
Db 347 LWMAPDGLYAKRLCQSRIYWD--GFLALCNDRENK---LERDQCKLFDQQLAELQAF 400
Qy 417 FRARQRRGSPRYTYILGFGQD 437
Db 401 F-AHGRSLPRFQVTLFCGEE 420

RESULT 12
Q90W10
ID Q90W10 PRELIMINARY; PRT; 416 AA.
AC Q90W10
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Interferon regulatory factor 10.
GN IRF10.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RA Nehyba J., Hrdlickova R., Bose H.R. Jr.;
RT "Interferon regulatory factor 10, a novel family member. IRF-10
RT expression is induced by interferons, concanavalin A, and the Rel/NF-
RT kappaB oncoprotein v-Rel.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF380350; AAK55444.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 416 AA; 47646 MW; 669967280FEA967C CRC64;

Query Match 11.8%; Score 321; DB 13; Length 416;
Best Local Similarity 25.9%; Pred. No. 1.7e-14;
Matches 126; Conservative 65; Mismatches 161; Indels 134; Gaps 27;

Qy 16 EWLIGISSCYEGLOWLDEARTCFRPWKHFARKDL-SEADARIFKAWAVARGRPPSS 74
Db 12 EWLIAQIDSGRYPLGWENRERLFRIPWKHAAQDYRQQDAALFRAWAVYKGYHEGT 71
Qy 75 RGGGPPPEATBAGKWNFRCALSTRFRFVMLRNSG-DPADPHKVVALSRELWCWG 133
Db 72 -----DKADPSTWTKRLCALNKSTDFQVPSRQIDISEPKYQI---VC--DG 117
Qy 134 PGTDQTEAEAPAAVPPPGGPGPL-AHTHAGLQAPGLPAPAGDKGDLILLQAVQOSCLA 193
Db 118 TRDAEKDEKGRMQPTSSKDPOGHVAESHG-----TAG-----TC-- 154
Qy 194 DHLITASWGADPVPTKAPGEGQELPLTGACAGGPGLPAGELYGWAVETTPSPGPQPAAL 253
Db 155 -HMPTL-----PLTAAHPAE--RGYHVRGI-----FYGM-----SP----- 182
Qy 254 TTGAAAAPSPHQAEPYLSFSPSACTAVQBPFGALDVTIMY-----KGRTV 300
Db 183 -TRSHLLPRAP-SFLPAEDVNHSDCW-----LHIRLYCDVLVKELTTRTAGCRI 231
Qy 301 LQKVVG--HPSCTFLYGPDPVAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLH 358
Db 232 ASRTEGYPGSCM-----EQIEFPPPALGGGWTAVTVEVLERLLP--H 274
```

Qy	359	LHLRGPQLWA	-----RRMKCKVYWEVGGP--PGSASPSTPACLLPRNCDTPIFDRV	400
Db	275	LE-RGVLLHWAPG	VFMKRCQCGRYVW--NGSLAPHQDWPNK-----LREKTYKLLDTQQ	327
Qy	410	FFOEIVFRAQRRC	SPYTYLGFQOD--LSAGRPKESLVLVKLEPWLCRVHLEGTOR	467
Db	328	FLELRRYLS-HGQ	PAPOYIHLCEGEYPTSTGRHLOK-LIMAHVEPVFARLFFHQAOR	385
Qy	468	EGVSSL	473	
Db	386	LGPALL	391	
RESULT 13				
Q9QZL7		PRELIMINARY;	PRT; 330 AA.	
ID	Q9QZL7			
AC	Q9QZL7			
DT	01-MAY-2000	(TRENBLrel. 13, Created)		
DT	01-MAY-2000	(TRENBLrel. 13, Last sequence update)		
DT	01-JUN-2001	(TRENBLrel. 17, Last annotation update)		
DE	IRf6	(Fragment).		
GN	IRF6.			
OS	Mus musculus	(Mouse).		
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria;	Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/J;			
RA	Sands A., Mak T.W.;			
RL	Submitted (AUG-1999)	to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF177668;	AAFO0915.1; -.		
DR	HSSP; P23906;	2IRF.		
DR	InterPro; IPR001346;	IRF.		
DR	Pfam; PF00605;	IRF. 1		
DR	PRINTS; PR00267;	INTFRNREGFCT.		
DR	PRODOM; PD002355;	IRF; 1.		
DR	SMART; SMO0348;	IRF; 1.		
DR	PROSITE; PS00601;	IRF; 1.		
FT	NON TER	1		
FT	NON TER	330 330		
SQ	SEQUENCE	330 AA; 37355 MW; EB02EC8B751CBD7D CRC64;		
Query Match 10.3%; Score 281.5; DB 11; Length 330;				
Best Local Similarity 26.0%; Pred. No. 6.9e-12;				
Matches 103; Conservative 49; Mismatches 161; Indels 83; Gaps 15				
Qy	29	GLQWLDEARTCFRVPWKIFAR-KDLSADARIFKAWAVARGWPPSSRGGGPPPEAE	TAE	87
Db	2	GLIWLHRDSKRFQIPWKIATRHSPQEEENTIFKAWAVETGKY---QBGVD	DDPAK---	55
Qy	88	RAGKNTFRCAIRLSTRFRVLMRLNSGD--PADPHKVYALSRELWCWEGPCTDQTEAE	APAA	146
Db	56	---WKAQRCAINLSREFNLMTDGTKEVPMNPVKIY---QVC-----		92
Qy	147	VPPQGGPPGPFLLARHATHAGLOAPGALPAPAGDKGLLLQAVQOSCLADHLLTAS	WGADPV	206
Db	93	IPQTQ-----SVINPGSTGSPWDEKNDVDEDEE-----EDEL	EQSQHHVPI	136
Qy	207	PTKAPGEGQGLPL-----TGACAGGPGPLPAGELYGWAVETTPSPGQP	PAALTGTGEAAPE	262
Db	137	QDTFFPLNINGSMPAPASVGNCSVGNCSF-----ESVWPKEPTEMEVPQAPI		184
Qy	263	SPHOAEPLVSPSPSACTAVQSPGALDVTIMYKGRTVLQKV-VGHP-SCTFLY	GPPDP-	319
Db	185	QPFYSSPELWISSPMT-----DLDIKFQYRKGYKGTWTVSNPQGR	LFYDGLGFM	236
Qy	320	-----AVRATPDQVAFPSPAELPDQQLRYTEELLRHVAPGLHLELRGPOL	WARRMKGC	374
Db	237	PDQBELGCPVSLQVKVFCPEHITNEKQKLFISKLLDVMDRGLLILEVSGHAI	YATRLCQC	296
Qy	375	KVYWEVGGPPGSPSTPACLLPRNCDTPIFDRV	F	410

RA Kinaishi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN Mol. Microbiol. 21:77-96(1996).
RP [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL133220; CAB61705.1; -
DR InterPro; IPR005158; BAD.
DR InterPro; IPR000767; Disease resist.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF03704; BAD; 1.
DR PRINTS; PR00364; DISEASERSIST.
DR PRINTS; PR01574; TUBBYPROTEIN.
DR ProDom; PD000329; Trans reg C; 1.
SQ SEQUENCE 1334 AA; 138787 MW; 78DC746883E8778C CRC64;

Query Match 6.7%; Score 183; DB 16; Length 1334;
Best Local Similarity 28.1%; Pred. No. 0.00021;
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;

Qy 18 LLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIKAWAVAGRWPPSSRGG 77
Db 146 VLADLPDRTAEARW--ETRHFEALRARHTAALDLGOAHSLELTALCDGH----- 195

Qy 78 GPPPEAETABR-----AGWKTNFRCALRSTRFRVLRDMSG-DPA-----D 117
Db 196 -PLDEPLQALRLRALRDSGRTEALAAVEAVRR--LLADRLGTDPGPELTHAELLSPS 252

Qy 118 PHKYALSRLELCWREGP-----GTDQTEAEAPAAVPPQGGP-PGPFLLANTHAGL 166
Db 253 FTPTPGRSRTPGWTSGPGPASGAGAASGTDVASCAGAAAGDPDPASGPASGPAVAPGSGG 312

Qy 167 QAPGLPAPAGDKDLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGGGLPLTGACAG 226
Db 313 PAFGWWPAPGTAPGSSTAPPHDTASAAD---TA---PAPGTSAPGTA---PAAGTAAP 362

Qy 227 GPGL--PA-GELY--GWA--VETTPSGPQFAALTGCEA-----AAPESPHOAEFYL 271
Db 363 APGTAGPAGTSYAPGTAPVAGTTPAGTAPAGTAPAGTAPAGTAPAGTAPVAGTTPAGT 422

Qy 272 SPSPSACTAV-----QEPSGALDVTIMYKRTVLQKVGHPSCTFLYGPDPDAVRAT 324
Db 423 APAPGSTPAPGTVPAGTAPAGPQPA-----DGR---RPVTGPASGT---GPG---AAT 468

Qy 325 DPQOVA-----FPSPA 335
Db 469 PPEAAAAAASAGSPSPA 485

Search completed: June 18, 2003, 12:47:49
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:45:44 ; Search time 180 Seconds
(without alignments)
1430.977 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAAPRVLFGEWLLG.....SANSLYDDIECFMLELEQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2715	99.4	503	1	IRF7_HUMAN	Q92985 homo sapien
2	1562	57.2	457	1	IRF7_MOUSE	P70434 mus musculus
3	1116	40.9	210	2	Q9UE79	Q9ue79 homo sapien
4	780	28.6	491	1	IRF3_CHICK	Q90643 gallus gall
5	778	28.5	491	2	Q90ZD4	Q90zd4 gallus gall
6	603	22.1	128	2	Q9HB64	Q9hb64 homo sapien
7	520	19.0	423	2	Q6PE40	Q6pe40 brachydanio
8	515.5	18.9	421	2	Q802H9	Q802h9 carassius a
9	487.5	17.9	498	1	IRF5_HUMAN	Q13568 homo sapien
10	484.5	17.7	488	2	Q6RCM8	Q6rcm8 homo sapien
11	482.5	17.7	514	2	Q64GB2	Q64gb2 homo sapien
12	458	16.8	427	1	IRF3_HUMAN	Q14653 homo sapien
13	453.5	16.6	419	2	Q764M6	Q764m6 sus scrofa
14	448	16.4	466	2	Q72XG6	Q72xg6 xenopus lae
15	445.5	16.3	492	2	Q6PCZ7	Q6pcz7 brachydanio
16	444.5	16.3	497	1	IRF5_MOUSE	P56477 mus musculus
17	440	16.1	459	2	O5Y578	O5y578 xenopus lae
18	439.5	16.1	382	2	Q6NZY9	Q6nzy9 brachydanio
19	438	16.0	419	1	IRF3_MOUSE	P70671 mus musculus
20	436.5	16.0	460	2	Q6GR23	Q6gr23 xenopus lae
21	436	16.0	467	1	IRF6_MOUSE	P97431 mus musculus
22	431	15.8	467	1	IRF6_HUMAN	O14896 homo sapien
23	430	15.7	467	2	Q9N136	Q9n136 ovis aries
24	428	15.7	467	2	Q91VD0	Q91vd0 mus musculus
25	424	15.5	517	2	Q63ZP8	Q63zp8 xenopus lae
26	422.5	15.5	460	2	Q6DD26	Q6dd26 xenopus lae
27	420	15.4	467	2	Q8WNQ4	Q8wnq4 sus scrofa
28	398	14.6	413	2	Q64GA9	Q64ga9 homo sapien
29	376	13.8	450	1	IRF4_MOUSE	Q64287 mus musculus
30	368	13.5	423	2	Q6DHK7	Q6dhk7 brachydanio
31	364.5	13.3	452	2	Q96GL3	Q96gl3 homo sapien

32	364.5	13.3	452	2	Q7Z5G6	Q7z5g6 homo sapien
33	364	13.3	445	2	Q98TX7	Q98tx7 gallus gall
34	363	13.3	451	1	IRF4_HUMAN	Q15306 homo sapien
35	362	13.3	409	2	Q98TX6	Q98tx6 gallus gall
36	344	12.6	440	2	Q99419	Q99419 homo sapien
37	338	12.4	425	1	ICSB_CHICK	Q90871 gallus gall
38	333.5	12.2	424	1	ICSB_MOUSE	P23611 mus musculus
39	331	12.1	392	2	Q6NY63	Q6ny63 brachydanio
40	328.5	12.0	393	1	IRTF_HUMAN	Q00378 homo sapien
41	322	11.8	397	2	Q6NS16	Q6ns16 xenopus lae
42	321	11.8	416	2	O90W10	O90w10 gallus gall
43	318.5	11.7	389	2	O64GB0	O64gb0 homo sapien
44	308.5	11.3	426	1	ICSB_HUMAN	Q02556 homo sapien
45	306	11.2	398	2	Q68FP4	Q68fp4 rattus norv

ALIGNMENTS

RESULT 1
ID IRF7_HUMAN STANDARD; PRT: 503 AA.
AC Q92985; O00331; O00332; O00333; O75924;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Interferon regulatory factor 7 (IRF-7).
GN Name=IRF7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Splice;
RA Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,
RA Sutherland G.R., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX MEDLINE=97459673; PubMed=9315633;
RA Zhang L., Pagano J.S.;
RT "IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency.";
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM D).
RX MEDLINE=99003279; PubMed=9786932; DOI=10.1074/jbc.273.44.29210;
RA Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pitha P.M.;
RT "Characterization of the interferon regulatory factor-7 and its potential role in the transcription activation of interferon A genes.";
J. Biol. Chem. 273:29210-29217(1998).
RN [4]
RP SPLICED ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.E., Brenner S.E.;
RT "An unappreciated role for RNA surveillance.";
Genome Biol. 5:RSEARCH008.1-RESEARCH008.16(2004).
CC -!- FUNCTION: Transcriptional activator. Binds to the interferon-promoter (Op) of EBV nuclear antigen-1 (EBNA1).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=Q92985-1; Sequence=Displayed;
CC Name=B; Synonyms=Beta;
CC IsoId=Q92985-2; Sequence=VSP_002760;
CC Name=C; Synonyms=Gamma;
CC IsoId=Q92985-3; Sequence=VSP_002758, VSP_002759;
CC Notes=May be produced at very low levels due to a premature stop codon in the mRNA, leading to nonsense-mediated mRNA decay;

Qy 2 ALAPRAAPRVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIFK 61

```
Db 3 ALDSEGAQKLRGFWLLNAVSSGLYRGLCWIDPDRIFRIPWKHNARKDVTSSDVEIFK 62
QY 62 AWAVARGRWPPSGRGPPPEAEAEARAGWKTNFRCAIRSTRFRVLMRDNSSGDPADPHKV 121
Db 63 AWAKASGRY-----EGNAEDPAKWKTNFRCAIRSTRHFMFLEDRSKCNDPDHVK 111
QY 122 YALSRELWCWREGPGTQTEAEAPAAVPPQ--GGP-----PGPFLAHTHAGLQ---- 167
Db 112 YAVA-----SGVENDRGSGGPVAGALQQOQPOLLNHHDLALENTPT 152
QY 168 -ARGPLPAPAGDKGL-LLQAVQOOSCLADHLLTASWADPVPTKAPGEG--QSGPLPTG- 222
Db 153 DSTEGVAAAALTOVDLQSLVQHCHNISAL-----GSPQTLWAHTGDALPEDALLPGQ 207
QY 223 -ACAGGGLPAGLYGW-AVETTPSPGPQ-----AALTTCGAAAPSPHOAEP 269
Db 208 DGLCLPG-----QFQWRQLEBELLGNQPLTGGCGGQDAGALPVSECAIPAPSPAE 262
QY 270 YL-----SPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGHPSCTFLYGPDPVAVRATD 325
Db 263 LLFQSANPAPPPAGDIGGLPPLDITIYRGMVYQEQVDDSRCLVAYQLDPAV--AE 320
QY 326 PQQVAPSPAEPLDQKQRLRYTELLRHVAPGLHLELPGPOLWARRMKCKVYWEV----- 380
Db 321 QRLVLPSPASLPDPQRRTYENLLE--VAGLRLEORAGOLLATRLKCKVFWALSQOLE 378
QY 381 GGPFGSASPTACLLPRNCDTPIFDFRVFFQELVEFRARQRRGSPRYTYILGFGQDLSA 440
Db 379 GGEP-----PLNLLHRDQETIFDFRVFCTELDRFRDRRRERSPDFIFLCFGQCFSS 431
QY 441 GRPKESLVILVLEPWLRCVHLEGTQREGVSSLDSDLDLCLSSANSYLDIECFLEMELE 500
Db 432 TKPKESKILVLPVPCFYVYEQVQRGASSLNSGNVSLQSLDSFNLFLFIQYHMQTD 491
```

RESULT 5

```
Q90ZD4 PRELIMINARY; PRT; 491 AA.
ID Q90ZD4;
AC Q90ZD4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Interferon regulatory factor-3.
GN Name=IRF-3;
OS Gallus Gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA May D.L., Grant C.E., Deeley R.G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268079; AAK58583.1; -.
DR HSSP; P15314; 1IF1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; C:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTFRNREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 491 AA; 54440 MW; 90E0C80A2624B1B4 CRC64;
```

Query Match 28.5%; Score 778; DB 2; Length 491;
Best Local Similarity 38.5%; Pred. No. 2.2e-35;
Matches 208; Conservative

```
QY 2 ALAPERAAPVRLGEMWLGISGCTEGLQWLEARTCFRVPWKHFARKDLSADARIFK 61
Db 3 ALDSEGAQKLRGFWLLNAVSSGLYRGLCWIDPDRIFRIPWKHNARKDVTSSDVEIFK 62
```

RESULT 6

```
Q9HB64 PRELIMINARY; PRT; 128 AA.
ID Q9HB64;
AC Q9HB64;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Interferon regulatory factor-7H (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20493557; PubMed=10924517; DOI=10.1074/jbc.M005288200;
RA Lu R., Au W.-C., Yeow W.-S., Hageman N., Pitha P.M.;
RT "Regulation of the promoter activity of interferon regulatory factor-7
gene: ACTIVATION BY INTERFERON AND SILENCING BY HYPERMETHYLATION.";
RL J. Biol. Chem. 275:31805-31812(2000).
DR EMBL; AF271159; AAG30003.1; -.
DR HSSP; P23906; 2IRF.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003677; F:RNA binding; TAS.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0006954; P:inflammatory response; TAS.
DR GO; GO:0046731; P:passive viral induction of host immune resp. .; TAS.
DR GO; GO:0006974; P:response to DNA damage stimulus; NAS.
DR GO; GO:0006615; P:response to virus; TAS.
DR GO; GO:0006367; P:transcription initiation from Pol II promoter; TAS.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTFRNREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
DR NON_TER 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;
SQ SEQUENCE 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;
```

```

Query Match          22.1%; Score 603; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.8e-26;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVFGWLLGRISSCYEGLQWLDDEARTCFRVPWKHFARKDSEADARIFKAWAVA 66
Db 20 RAAPRVFGWLLGRISSCYEGLQWLDDEARTCFRVPWKHFARKDSEADARIFKAWAVA 79
QY 67 RGRWPPSSRGGGPPPEAETAEAGKTNFRCALRSTRFRVMLRDNSGDP 115
Db 80 RGRWPPSSRGGGPPPEAETAEAGKTNFRCALRSTRFRVMLRDNSGDP 128

RESULT 7
Q6PE40
ID AC Q6PB40 PRELIMINARY; PRT; 423 AA.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Interferon regulatory factor 7.
GN Name=irf7;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058298; AAH58298.1; -.
DR ZFIN; ZDB-GENE-040426-1518; irf7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD FHA.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTFRNREGFT.
DR ProDom; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
SQ SEQUENCE 423 AA; 49060 MW; 3A59DE6FB6D2275E CRC64;

Query Match          19.0%; Score 520; DB 2; Length 423;
Best Local Similarity 30.9%; Pred. No. 3.5e-21;
Matches 151; Conservative 67; Mismatches 174; Indels 96; Gaps 16;

QY 14 FGEWLLGRISSCYEGLOWLDEARTCFRVPWKHFARKDSEADARIFKAWAVARGR---W 70
Db 10 FGPWLLIEQVESQYEGLSMI--GHDIFRIPWKHNARRDLGDADVKIFKEWAIYVSGKINEY 67
QY 71 PPSRRGGGPPPEAETAEAGKTNFRCALRSTRFRVMLRDNSGDPADPHKVVALSRELWC 130
Db 68 P-----NDKAKWKNFRCALHSLKUNFEMLEHDSKDPDDQHKIYRIIR----- 109
QY 131 REGPGTDTAEAPAAVPPPGGPPG-PFLAHTHAGLAQAPLPAPAG-----DKGDL 182
Db 110 -----PQNHQEIQSAIQSAEAVQQLFFIAEVY---NASNHSQDMLELLNLVETMDL 160
QY 183 LQAVQOQSLADHLHTASGADPVPTKAPGEGQEGPLTGTACAGGPGLPAGBELYGHAVET 242
Db 161 NLHAVSQS-----LXTYS-----QPNIQTTSSN-----YFET 187
QY 243 TPSPGPQPAALTGGEAAAPESPQAEPLYSPPSACTAEOEPSGALDVTIMYKGRVTLO 302
Db 188 TVSDGP---CMQNNIPASVQOSH-----TTVDQWNLCDLEISINRTRTEVLK 231
QY 303 KVVGHPSCTFLYGPDPDPAVRATDPOQVAFPPSPAEPLDQKQLRYTBELLRHVAPGLHLELR 362
Db 232 TFLCSSLIHFFVQCDPSBLRG---EQIRFPTTECLIDVKIQIYTKRILDSIQRLQLEVN 288
QY 363 GPQLWARMGKCKVYVEGPGPSASPTPACLLPRNCDTPIDFRVFFQELVEFRAROR 422
Db 289 QYGIYGFQRDKCKVFVSTSDPCEIQKPEPRK--LQNVYKEQLLSFDKYIRDLDDFK-ENR 345
QY 423 RGSPTYTYLGFQDLSAGRPKEKSLVLVLEKLPWLCRVHLEGTORGESVLSDDLDLCL 482
Db 346 GGSPTYTYLGFQDLSAGRPKEKSLVLVLEKLPWLCRVHLEGTORGESVLSDDLDLCL 405
QY 483 SSANSLYD 490
Db 406 SH-NSLYD 412

RESULT 8
Q802H9
ID Q802H9 PRELIMINARY; PRT; 421 AA.
AC Q802H9;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Interferon regulatory factor 7.
GN Name=irf7;
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Zhang Q., Xu D., Hu C., Gui J.;
RT "Identification of antiviral-relevant genes in the cultured fish cells
RL induced by inactivated virus.";
RL Chin. Sci. Bull. 48:581-588(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22912602; PubMed=14550671; DOI=10.1016/S1050-4648(03)00025-1;
RA Zhang Y., Hu C., Zhang J., Huang G., Wei L., Zhang Q., Gui J.;
RT "Molecular cloning and characterization of crucian carp (Carassius
RL auratus L.) interferon regulatory factor 7(1).";
RL Fish Shellfish Immunol. 15:453-466(2003).
DR EMBL; AY177629; AAO18646.1; -.
DR HSSP; F15314; IIF1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
SQ SEQUENCE 423 AA; 49060 MW; 3A59DE6FB6D2275E CRC64;

```



```
QY 259 AAPESHQAEPLYSPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGH-P-SCFTFLYGP 317
DB 228 --PPAGEQLPDLISPHMLPLTD-----LEIKFOYRGRPPRALTISNPHGCRFLYS-- 277
QY 318 DPAVRATDQ-----OVAFPPSAELPDOKQLRYTEELLRHVAPGLHLELGPQLW 367
DB 278 --OLEATQOVELFGPISLEQVRFPSPEDIPSDKQRYTNQLLDVLDRLGLILQLOQGDLY 335
QY 368 ARMGCKVYWEVGGPGSASPTACLLPRNCDTPIFDPRVFFOBLVEFRARQRGSPR 427
DB 336 AIRLCOCKVFW--SGPCASAHDCPN-PIQREVKTGLFSLEHFLNELILFQKQTNTPPP 392
QY 428 YIYLFGQDLSAGRPKESILVVKLEPWLCRVHLEGTQREGVSSLDSDLDL 480
DB 393 FEIFFCGEWPDRKPREKLIITVQVVPVAARLLEMFSGELSWADSIRLQI 445

RESULT 10
Q6RCM8 PRELIMINARY; PRT; 488 AA.
AC Q6RCM8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Interferon regulatory factor 5 variant 4 (Interferon regulatory factor
DE 5 variant 3).
GN Name=IRP5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21316468; PubMed=11303025; DOI=10.1074/jbc.M101216200;
RA Barnes B.J., Moore P.A., Pitha P.M.;
RT "Virus-specific activation of a novel interferon regulatory factor,
RT IRF-5, results in the induction of distinct interferon alpha genes.";
RL J. Biol. Chem. 276:23382-23390(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Barnes B.J., Pitha P.M.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY504947; AAR90326.1; -.
DR EMBL; AY504946; AAR90325.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR009058; wing_hlx DNA_bnd.
DR Pfam; PF00605; IRF. 1.
DR PRINTS; PR00267; INTERFERGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 488 AA; 54942 MW; 4E43B54EE74A7E37 CRC64;
```

```
Query Match 17.7%; Score 484.5; DB 2; Length 488;
Best Local Similarity 28.6%; Pred. No. 3.6e-19;
Matches 151; Conservative 59; Mismatches 171; Indels 147; Gaps 20;
```

```
QY 1 MALAPERAAPRVLFGEWLGLGEISSCYEGLQWLDEARTCFRVPWKHFARKDLS-EADARI 59
DB 7 VAPTTPR---RVRLKFWLVAQVNSCYQPGLOWNGEKKLFCIPWRHATHRGPSQGDNTI 63
QY 60 FXAWAVARGWPPSSRGGGPPPEATAERAGWKTNFRCAALSTRFRVLMRLDSGD--PADP 118
DB 64 FXAWAKETGKYTEG-----VDEADPAKWANKALNCALNKSDFRLIYDGRDMPPOP 114
QY 119 HKVYALSRELCHWREGPGTDQTEAEAPAAVPPQGGPPG-----PFLAHTH 148
DB 115 YKIY-----EVC-SNGP-----APTDSQPPEDYSFGAGEEEEEELQRLMPLSLTSDYVKWPTTLQ 170
```

```
QY 149 PPOGPPP---GFLAHTHAGLQAPGFLPAPAGDKGDLLOAVQOQCLADHLLTASWGADP 205
DB 171 PPTLQPPVVVLG-----PAP-----DP 187
QY 206 VPTKAGBQEGGLPLTGACAGGGLPAG--ELYGWAVETTPSGPOPAAATTGEEAAAPES 263
DB 188 SPLAPP-----PGNPAGFRELLSEVLE-----PGPLPASL-----PPA 220
QY 264 PHQAEPLYSPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGH-P-SCFTFLYGPDPAPVR 322
DB 221 GFQLLPDLLISPHMLPLTD-----LEIKFOYRGRPPRALTISNPHGCRFLYS-----QLE 270
QY 323 ATDPO-----QVAFPPSAELPDOKQLRYTEELLRHVAPGLHLELGPOLWARRMG 372
DB 271 ATQEOVELFGPISLEQVRFPSPEDIPSDKQRYTNQLLDVLDRLGLILQLOQGDLYAIRLC 330
QY 373 KCKVYWEVGGPGSASPTACLLPRNCDTPIFDPRVFFOBLVEFRARQRGSPRTIYL 432
DB 331 QCKVFW--SGPCASAHDCPN-PIQREVKTGLFSLEHFLNELILFQKQTNTPPPPEIFF 387
QY 433 GFGQDLSAGRPKESILVVKLEPWLCRVHLEGTQREGVSSLDSDLDL 480
DB 388 CFGEWPDRKPREKLIITVQVVPVAARLLEMFSGELSWADSIRLQI 435

RESULT 11
Q64GB2 PRELIMINARY; PRT; 514 AA.
AC Q64GB2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Interferon regulatory factor 5 variant 5.
GN Name=IRP5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu G., Mancel M.E., Barnes B.J.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY693665; AAU12877.1; -.
SQ SEQUENCE 514 AA; 57770 MW; D554D121D9E35893 CRC64;
```

```
Query Match 17.7%; Score 482.5; DB 2; Length 514;
Best Local Similarity 29.6%; Pred. No. 4.9e-19;
Matches 154; Conservative 60; Mismatches 201; Indels 105; Gaps 19;
```

```
QY 1 MALAPERAAPRVLFGEWLGLGEISSCYEGLQWLDEARTCFRVPWKHFARKDLS-EADARI 59
DB 7 VAPTTPR---RVRLKFWLVAQVNSCYQPGLOWNGEKKLFCIPWRHATHRGPSQGDNTI 63
QY 60 FXAWAVARGWPPSSRGGGPPPEATAERAGWKTNFRCAALSTRFRVLMRLDSGD--PADP 118
DB 64 FXAWAKETGKYTEG-----VDEADPAKWANKALNCALNKSDFRLIYDGRDMPPOP 114
QY 119 HKVYALSRELCHWREGPGTDQTEAEAPAAVPPQGGPPG-----PFLAHTH 163
DB 115 YKIY-----EVC-SNGP-----APTDSQPPEDYSFGAGEEEEEELQRLMPLSLTSD 161
QY 164 AGLQAPGPLP-----APAGDKGDLLOAVQOQCLADHLLTASWGADPVPTKAPGE 213
DB 162 AVQSGPHMTFYSLLKEDVKWPTTLQPTTLRPTTLQPTLPVVLGPPADPSPLAPP-- 219
QY 214 GQEGPLTGACAGGGLPAG--ELYGWAVETTPSGPOPAAATTGEEAAAPESHPQAEPLY 271
DB 220 -----PCNPAGFRELLSEVLE-----PGPLPASL-----PPAGEQLLPDL 254
QY 272 SPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGH-P-SCFTFLYGPDPAPVRATDQ--- 327
DB 255 LISPHMLPLTD-----LEIKFOYRGRPPRALTISNPHGCRFLYS-----OLEATQOVEL 304
```

328 -----QVAFPPSPAEIPDQKQRYTELLRHVAPGLHLELRGPOLWARMGKCKVYEV 380
305 FGPISLEQVFPSPEDIPSKQRYTNQLDLVDRLGLLQGLQDLYAIRLCKCKVFW-- 362
381 GGPFGSASPSTACILPRNCDTIFDPRVFQELVEFRARQGRGSPRYTYILGFGQDLISA 440
363 SGPCASAHDCPN-PIQREVKTKLFLSLHFLNELILFKQGTNTPTPPFFFCGSEWPD 421
441 GRPKESLVLKLEPMLCRVHLEGTQREGVSSLDL 480
422 RKPREKKLITVQVPPVAARLLLEMFSGELSWADSIRLQI 461
RESULT 12
IRF3_HUMAN STANDARD; PRT; 427 AA.
AC Q14653;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Interferon regulatory factor 3 (IRF-3).
GN Names:IRF3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Au W.-C., Moore P.A., Lowther W.W., Juang Y.-T., Pitha P.M.;
RT "Identification of a member of the interferon regulatory factor family
RT that binds to the interferon-stimulated response element and activates
RT expression of interferon-induced genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:11657-11661(1995).
RN [2]
RP SEQUENCE OF 323-413 FROM N.A.
RX MEDLINE=99020108; PubMed=9803267;
RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.;
RT "Mapping of a human interferon regulatory factor 3 (IRF3) to chromosome
RL 19q13.3-13.4 by an intergenic polymorphic marker.";
RL Ann. Hum. Genet. 62:231-234(1998).
CC -I- FUNCTION: May activate transcription by complex formation with
CC other transcription factors, possibly members of the STAT
CC family. Binds specifically to the IFN-stimulated response element
CC (ISRE) but not to the IRF-1 binding site PRD-I.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- TISSUE SPECIFICITY: Expressed constitutively in a variety of
CC tissues.
CC -I- SIMILARITY: Belongs to the IRF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z56281; CA91227.1; -
DR EMBL; U86636; AAC6818.1; -
DR PDB; 1J2F; X-ray; A/B=170-427.
DR PDB; 1QWT; X-ray; A/B=173-427.
DR TRANSFAC; T04673; -
DR Genew; HGNC:6119; IRF3.
DR H-InvDB; HIX0015338; -
DR MIM; 603734; -
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0003712; P:transcription cofactor activity; TAS.
DR GO; GO:0003666; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.

DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF. 1.
DR PRINTS; PR00267; INTERNRREGPCT.
DR ProDom; PD002355; IRF. 1.
DR SMART; SM00348; IRF. 1.
DR PROSITE; PS00601; IRF. 1.
KW 3D-structure; Activator; DNA-binding; Nuclear protein; Polymorphism;
KW Transcription regulation.
FT DNA BIND 7 107
FT VARIANT 96 96
FT VARIANT 377 377
FT VARIANT 427 427
FT VARIANT 427 427
SQ SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;
Query Match 16.8%; Score 458; DB 1; Length 427;
Best Local Similarity 28.0%; Pred. No. 9.3e-18;
Matches 141; Conservative 64; Mismatches 191; Indels 108; Gaps 19;
QY 10 PRVLFGELLGEISSCYEGLOWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVARGR 69
DB 6 PRIL--PWLVSQDLGLQGVAVWVNSRTRFRIPWKHGLRQDAQOQDEFGIFQAWATGA 63
QY 70 WPPSSRGSGGPPPEAETAEAGWKTFRCALRSTRFRVLMRLDNSGDPADPHKVYALSREL 129
DB 64 YVP---GRDKPDLPT---WKRNFRSALNRKEGLAEDRSKDPHPHKIYEFV--- 110
QY 130 WREGPTDTEAEAPAAVPPPGGPGPFLAHTHAGLQAPGLPAPAGDKGDLQLLQAVQQ 189
DB 111 -NSGVG---DFSQPDTSPTDNGG-----GSTSD-----TQ 136
QY 190 SCLADHLITASGADVPYKAPGEGQGLPLTCACAGGGLPAGELYGWAVETTPSPGPQ 249
DB 137 EDILDELL--GNMVLAPLP-----DFGPP-----SLAVAPFPCPQ 169
QY 250 PAALTGEEAAAPESPHQABPYLSPSPSACTAVOEP--SPGALDVTIMYKGRITVLQVVGHP 308
DB 170 PLRSPSLDNTPT-----FNLGSPENPLKRLVPGEEFEVTAFYRGQVQQTISCP 223
QY 309 SCTFLYGPDDPAVRATDPOQVAPSPA-ELPDQKQLRYTEELLRHVAPGLHLELRPQLW 367
DB 224 EGLRLVG-SEVGDRTLPFGWPVTLDPFGMSILTRGVMSYVRHVLSCGGGLALWRAGQLW 282
QY 368 ARWGMCKVYWEVG-----GPPGSASPSPACILPRNCDTPIFDRVFQELVEPR 418
DB 283 AQLRGHCHTYWAVSEELLNSGHGPDGE-----VPAKKEGVDFDLGPFIVDLITFT 333
QY 419 ARQRRGSPRYTYILGFGQDLISAGRPKEKSLVLYKLEPMLCRVHLEGTQREGVSSLDSSDL 478
DB 334 EGSGR--SPRYALWFCVYGESWPDQWPWKRLVMVYVPTCLRALVEMARVGASSLENT-V 391
QY 479 DLCLLSANSL---YDDIEFLMEL 499
DB 392 DLHISNSHPLSLTSDQYKAYLQDL 415
RESULT 13
Q764M6 PRELIMINARY; PRT; 419 AA.
ID Q764M6
AC Q764M6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Interferon regulatory factor 3.
GN Name:irf3;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.

[illegible]

Db 207 TPAYSTAPVHGAHAELP-----AEVATNOGSPLOQHITDHPQDRTL 251
QY 287 GALDVTIMYKGRVLOKVGHPSCFTLYGPPDPAVRATDP---QQVAPSPAEPLDQK 342
Db 252 TEFEVTVYRGIEVSKLVNKP-----HGRPRITSRHSSPGSVLDVLPSPMTIADQAV 306
QY 343 LRYTEELLRHVAPGLHLELRGPOLWARRMGKCKVYVEVGPPGSGASPTSPACILPRNCDT 402
Db 307 VGEIHKLLRNLEBGLVEVRGSGICGKQKQKCFRFSMTPEISQPNQ-----IDKNDYC 362
QY 403 PIFDFRVFOELVEFRARORRGSPRTYLYLGFQDLSAGRPKESLVLVKLEPWLRCVHL 462
Db 363 ILHTLQOFAELTEFIERTKSPQYHIMWCLGELPDRVPMNKKFMQVIVPSMKLLH 422
QY 463 EGTQRGVSLDSDLDLCLSSANSYLDIECEFLMELEQ 501
Db 423 DMSYSTGASSLHSSSEINLEISDSLSSTNDVMVLRLEHE 461

RESULT 15
Q6PGZ7 PRELIMINARY; PRT; 492 AA.

AC Q6PGZ7
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:63500.
GN ORFNames=zgc:63500;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerk A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC056772; AAH56772.1;
DR ZFIN; ZDB-GENE-040426-1137; zgc:63500.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTRNRREGFC.

DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Hypothetical protein.
SQ SEQUENCE 492 AA, 54876 MW, 98BF8FD3FE1F589B CRC64;

Query Match 16.3%; Score 445.5; DB 2; Length 492;
Best Local Similarity 27.4%; Pred. No. 5.3e-17;
Matches 147; Conservative 74; Mismatches 187; Indels 129; Gaps 25;

QY 11 RVLFGEWLGESGCGYEGLOWLDEARTCFRVPWKHFAR-KULSEADARLEKAWAVARGR 69
Db 7 RVRLKFWLVSQVDNATFFGLVWLDRDAKRFQPMGHATRHTPQOESEENTLEKAWAVETGK 66
QY 70 WPSRSSGGGPPPEAEATAERAGWKTNFRCAIIRSRFRVLMRLDNGSD-PADPHKYVYALSREL 128
Db 67 Y---QEGVDEPDPAK-----WKAQLRCALNKSREENLYDGTKEVPMPLKIY----DV 113
QY 129 CWREGFGTDTQTEAEAPAAVP-----PPQGGPPGPFLLAHTHAGLQAPGLPAPAG 177
Db 114 C-----DIPQTPSN-PGSPVPTVETDGEDDIPDTPEPPPPYASH---GMSA-----SP-- 157
QY 178 DKGDLLOAVQOSCLADHLLITASWGADPVTKAPGEGQGLPLTGACAGGPGLPAGELYG 237
Db 158 -----LAAVM-----TPGGSVSLQP--SSCAPPPALP----- 183
QY 238 WAVETTPSPG-POPAALTTGEAAAP-----ESPHQAEPLYLSPSPSACTAVQEPSPGAL 289
Db 184 -----PPALPPPAATVLPAAAPAPANVWPKEPEDVE--MQPPP-----MEIQPTAL 230
QY 290 D-----VTIMYKGRVLOKV-VGHP-SCFTLYGPPDPAVRATD--- 325
Db 231 DNLFITPWTWISSLPMTDLEVQFYRGKEVCPLTVSNPQGCRLFYDGLGPIVNBELFG 290
QY 326 ----POQVAFSPAEPLDQKQLRYTEELLRHVAPGLHLELRGPOLWARRMGKCKVYVEVG 382
Db 291 PVSLEQVRFPPTTEHIANDKQVFTSLLDVMGRGLILEVSGHDIYAVRLCQCKVYW---S 347
QY 383 PGASASPTPACILPRNCDTPIDFRVFFQELVEFRARORRGSPRTYLYLGFQDLSAGR 442
Db 348 GPCAPNPAPN-LIERQOKVKVLCLESFLSGVITHQRGSSDAPLDIHLFCFGEWPDGR 406
QY 443 PKEKSLVLKLEPWLRCRVHLEGTQREGVSSLDSSDLCLSSANSYLDIECEFLMEL 499
Db 407 PRERKLIMVQVVPVVARMISEMPSGDCSTRFSDSGSVRLQISIPD-IKDNIVTHLKL 462

Search completed: May 6, 2005, 16:53:59
Job time : 183 secs